A CHAOS THEORY AND NONLINEAR DYNAMICS APPROACH TO THE ANALYSIS OF FINANCIAL SERIES : A COMPARATIVE STUDY OF ATHENS AND LONDON STOCK MARKETS

A Dissertation

Presented for the

Doctor of Philosophy Degree

Warwick Business School

The University of Warwick, UK.

Aristotle D. Karytinos

June, 1999

A CHAOS THEORY AND NONLINEAR DYNAMICS APPROACH TO THE ANALYSIS OF FINANCIAL SERIES : A COMPARATIVE STUDY OF ATHENS AND LONDON STOCK MARKETS

A Dissertation

Presented for the

Doctor of Philosophy Degree

Warwick Business School

The University of Warwick, UK.

Aristotle D. Karytinos

June, 1999

DEDICATION

To my wife Katia

for her support and patience

TABLE OF CONTENTS

INTRODUCTION **1**

1. THE THEORETICAL FRAMEWORK OF CHAOS THEORY **4**

1.1. INTRODUCTION 4

1.2. DYNAMICAL SYSTEMS AND CHAOS DEFINITION 6

1.3. ATTRACTORS 10

1.4 DIMENSIONS 12

1.4.1. The Fractal Dimension *12*

1.4.2. The Correlation Dimension *14*

1.5. LYAPUNOV EXPONENTS 16

1.6. PHASE SPACE RECONSTRUCTION METHODS 19

1.6.1. The Method of Delays *20*

1.7. THE SINGULAR VALUE DECOMPOSITION METHOD 26

1.7.1. The SVD Analysis in noisy Cases and Noise Filtering *28*

2. THE CHAOTIC TESTING FRAMEWORK **31**

2.1 INTRODUCTION 31

2.2 THE CORRELATION DIMENSION ESTIMATION 32

2.2.1. Limitations and Pitfalls *35*

2.2.2. The Surrogate Data Method and the Bootstrap *41*

2.3 THE LARGEST LYAPUNOV EXPONENT ESTIMATION 47

2.4 ADDITIONAL METHODS TO IDENTIFY CHAOS 51

2.4.1 The "Signal Differentiation" Method *51*

2.4.2 The "Independent Realisations" Method *51*

3. NONLINEAR DYNAMICS AND CHAOS IN THE MARKETS **53**

3.1 A THEORETICAL PERSPECTIVE 53

3.1.1. The traditional (stochastic) framework *53*

3.1.2. The Chaotic alternative *55*

3.2 EMPIRICAL APPLICATION IN ECONOMICS 59

3.3 DATA AND METHODOLOGY 66

4. TESTS FOR INDEPENDENCE AND FRACTALITY....................................70

4.1 STATISTICAL DESCRIPTION....................................70

4.2 THE BDS TEST....................................75

4.2.1. Description of the test....................................75

4.2.2. Empirical evidence....................................77

4.2.3. The Nonstationarity issue....................................80

4.3 THE R/S ANALYSIS....................................83

4.3.1. Description of the method....................................83

4.3.2. Application problems....................................87

4.3.3. Empirical evidence....................................93

5. APPLICATION OF THE CHAOTIC TECHNIQUES....................................107

5.1 VISUAL INSPECTION TECHNIQUES....................................107

5.2 THE CORRELATION DIMENSION ESTIMATION....................................113

5.2.1. The ASE returns....................................113

5.2.2. The LSE returns....................................120

5.3 THE LARGEST LYAPUNOV EXPONENT ESTIMATION....................................124

6. NOISE FILTERING AND TESTING OF NOISE FILTERED SERIES......129

6.1 NOISE FILTERING BY THE SVD METHOD....................................129

6.1.1. Overview and implementation issues....................................129

6.1.2. Empirical Evidence....................................131

6.2 TESTING THE NOISE FILTERED SERIES....................................145

7. NONLINEAR PREDICTION METHODS AS DETECTION TOOLS FOR CHAOTIC DYNAMICS....................................153

7.1 INTRODUCTION....................................153

7.2 DESCRIPTION OF THE METHODS....................................156

7.2.1. The piecewise linear approximation method....................................156

7.2.2. The Simplex method....................................159

7.3 EMPIRICAL EVIDENCE....................................161

7.3.1. The DVS plot *260*

7.3.2. The varying prediction time and the dimensionality techniques *168*

7.4 CONCLUSIONS 176

8. SHORT TERM FORECASTING

AND ECONOMIC ASSESSMENT OF FORECASTS 177

8.1 INTRODUCTION 177

8.2 DESCRIPTION OF METHODOLOGY 179

8.3 EMPIRICAL EVIDENCE 186

8.3.1.ASE & LSE prediction results *186*

8.3.2.ASE & LSE prediction results (Sub-period analysis) *204*

8.4 CONCLUSIONS 233

9. SUMMARY OF RESULTS AND FINAL CONCLUSIONS 236

REFERENCES* *244

APPENDIX* *269

LIST OF ABBREVIATIONS* *273

ACKNOWLEDGEMENTS

There are a few people to whom I feel indebted for their contribution to the completion of this thesis.

My supervisor Professor Stewart Hodges, who supported my decision to work in this new research field despite the "risks" involved and who was always there when I needed his advice. Our discussions and his valuable comments have been decisive for the preparation of this study.

Professor Tassos Bountis (Dept. of Mathematics, University of Patras), for his useful comments on various technical issues related to chaotic techniques and methods. George Papaioannou, researcher at the University of Patras, for his valuable help in the software construction and our discussions related to the testing framework of this work.

My friend and ex colleague Panayiotis Iatrou, General Manager of "Delos" mutual funds, for illuminating discussions on the structure and dynamics of the Athens Stock Market. Professor Michael Glezakos (Dept. of Finance, University of Piraeus), who encouraged me in undertaking this endeavor. Last but not least, Ms. Theresa Messaris, a helping hand in editing this volume.

Finally, I would like to thank my family for their moral support and patience during my "absence" all these years.

DECLARATION

During the preparation of this thesis in the period from September 1992 to May 1997, a number of research and conference papers were prepared, as listed below. The remaining parts of the thesis have not been published.

[1] "Nonlinear Time Series Analysis of the Stock Exchange: The Case of an Emerging Market" published in the *International Journal of Bifurcation and Chaos*, Vol. 5 (6), 1995, written jointly with G. Papaioannou. The paper includes part of the analysis presented in Chapters 4, 5 and 6 and I hold full responsibility for the ideas put forward and the empirical implementation related to the contents of the above-stated Chapters.

[2] "Long-Term Dependence In Exchange-Rates", published in the *Journal of Discrete Dynamics in Nature and Society,* Vol 3 (1), 1999, written jointly with A. Andreou and G. Pavlides. The paper includes aspects covered in Chapter 4 of the thesis, applied to a different set of data provided by the co-authors. In this case too, I hold full responsibility for the ideas put forward and the empirical implementation related to the contents of Chapter 4.

[3] Parts from Chapters 4,5 and 6, were presented in a lecture titled: "Empirical findings from the nonlinear and chaotic analysis of the Athens Stock Market" during the "8th Annual International Conference and Workshops on Complex Dynamics and Chaos" held in Xanthi, Greece (July, 1995).

[4] The contents of Chapter 7 were presented in a lecture titled: "Nonlinear Forecasting Techniques as Tools for Chaos Identification" during the "9th Annual International Conference and Workshops on Complex Dynamics and Chaos" held in Xanthi, Greece (August, 1996).

[5] The contents of Chapter 8 were presented in a lecture titled: "Economic Assessment of Short-term Forecasts in Financial Markets: A case study of the

Athens and London Stock Exchange" during the "10th Annual International Conference and Workshops on Complex Dynamics and Chaos" held in Salonica, Greece (July, 1997).

[6] The contents of Chapters 2 and 3 were presented in a lecture titled: "Nonlinear Dynamics and Chaos: Applications in Economics and Finance" during the "11th Annual International Conference and Workshops on Complex Dynamics and Chaos" held in Livadia, Greece (July, 1998).

[7] "Non-Linear Time Series Analysis of the Greek Exchange-Rate Market", accepted for publication by *"The International Journal of Bifurcation and Chaos"*. The paper includes aspects covered in Chapter 4, 5 and 6 of the thesis applied to a different set of data.

[8] "Non-Linear forecasting of Financial Time Series: A comparative analysis", paper under preparation. The paper includes aspects covered in Chapter 8 of the thesis.

LIST OF TABLES

Table 3.1 Empirical literature – Summary of published results.......................... 60

Table 4.1 Statistical description of the daily returns of the ASE and LSE markets.. 70

Table 4.2 The BDS test (W statistic) for the ASE returns, filtered returns, scrambled filtered returns and a Gaussian random surrogate................ 78

Table 4.3 The BDS test (W statistic) for the LSE returns, filtered returns and scrambled filtered returns.. 79

Table 4.4 The BDS test (W statistic) for 3 consecutive subperiods of the ASE daily returns... 81

Table 4.5 The BDS test (W statistic) for 6 consecutive subperiods of the LSE daily returns... 82

Table 4.6 Asymptotic critical values of the modified R/S statistic.......................... 90

Table 4.7 Hurst estimates test of significance against two random alternatives of the daily ASE returns.. 95

Table 4.8 The modified R/S statistic of the ASE return series & bootstrapped critical values of the test statistic... 96

Table 4.9 Hurst estimates and test of significance against two random alternatives of the five-day ASE returns... 100

Table 4.10 Hurst estimates and test of significance against two random alternatives of the twenty-day ASE returns.. 101

Table 4.11 Hurst estimates and test of significance against two random alternatives of the daily LSE returns.. 104

Table 4.12 The modified R/S statistic of the LSE daily return series & bootstrapped critical values of the test statistic... 104

Table 4.13 Hurst estimates and test of significance against two random alternatives of the AR(9) filtered one-day returns, the AR(1) filtered twenty-day returns and the AR(1) filtered one-day returns........................... 105

Table 5.1 ASE: Statistical assessment of the "shuffle test" results.......................... 115

Table 5.2 ASE: Statistical assessment of the "randomised phase" test results......... 117

Table 5.3 ASE : Statistical assessment of the "randomised sign" test results.......... 118

Table 5.4 LSE: Statistical assessment of the "shuffle test" results.......................... 120

Table 5.5 LLE estimation for the ASE returns, (t_s = 15 and m = 6, 9, 13)............... 126

Table 5.6 LLE for different chaotic, random and periodic specifications 126

Table 5.7 LLE for the ASE and LSE original and shuffled series.......................... 127

Table 6.1 Noise filtered ASE series : Statistical assessment of the randomized phase test.......................... 147

Table 6.2 The modified R/S statistic for the ASE noise filtered series & bootstrapped critical values of the test statistic.......................... 150

Table 6.3 The modified R/S statistic for the LSE noise filtered series & bootstrapped critical values of the test statistic.......................... 150

Table 6.4 The modified R/S statistic for the "short" LSE noise filtered series & bootstrapped critical values of the test statistic.......................... 151

Table 6.5 LLE estimation for the ASE noise-filtered returns.......................... 152

Table 6.6 LLE for the ASE and LSE noise- filtered series.......................... 152

Table 8.1 ASE prediction results (Library: 1 year, Prediction period: 12 years)..... 188

Table 8.2 ASE prediction results (Library: 2 years, Prediction period: 11 years)... 189

Table 8.3 ASE prediction results (Library: 3 years, Prediction period: 10 years)... 190

Table 8.4 ASE prediction results (Library: 4 years, Prediction period: 9 years)..... 191

Table 8.5 ASE prediction results (Library: 5 years, Prediction period: 8 years)..... 192

Table 8.6 LSE prediction results (Library: 1 year, Prediction period: 12 years)..... 195

Table 8.7 LSE prediction results (Library: 2 years, Prediction period: 11 years)... 196

Table 8.8 LSE prediction results (Library: 3 years, Prediction period: 10 years)... 197

Table 8.9 LSE prediction results (Library: 4 years, Prediction period: 9 years)..... 198

Table 8.10 LSE prediction results (Library: 5 years, Prediction period: 8 years).... 199

Table 8.11 Comparison of prediction results of the PW model (LSE series) for shorter and longer library length and fixed prediction period.......... 201

Table 8.12 Annualised volatility and sign change frequency estimation for the ASE and the LSE series.......................... 206

Table 8.13 ASE prediction results (Highest volatility and Lowest Sign Change Frequency period/Library: 2 years, Prediction period: 1 year)............... 209

Table 8.14 ASE prediction results (Lowest volatility period/Library: 2 years, Prediction period: 1 year).......................... 210

Table 8.15 ASE prediction results (Highest sign change frequency period Library: 2 years, Prediction period: 1 year)....211

Table 8.16 Prediction improvement (%) of the Lowest vs. Highest volatility period, and Lowest vs. Highest Sign Change frequency period for the ASE returns....213

Table 8.17 ASE: Sub-period prediction results....216

Table 8.18 ASE: Spearman rank correlation analysis between prediction indicators and the criteria of volatility and sign change frequency....218

Table 8.19 ASE: Spearman rank correlation analysis between forecast accuracy measures and indicators related to economic results....219

Table 8.20 LSE prediction results (Highest volatility period/Library: 2 years, Prediction period: 1 year)....224

Table 8.21 LSE prediction results (Lowest volatility period/Library: 2 years, Prediction period: 1 year)....225

Table 8.22 LSE prediction results (Highest Sign Change Frequency period Library: 2 years, Prediction period: 1 year)....226

Table 8.23 LSE prediction results (Lowest Sign Change Frequency period Library: 2 years, Prediction period: 1 year)....227

Table 8.24 Prediction improvement (%) of the Lowest vs. Highest volatility period, and Lowest vs. Highest Sign Change frequency period for the LSE returns....228

Table 8.25 LSE: Sub-period prediction results....230

Table 8.26 LSE: Spearman rank correlation analysis between prediction indicators and the criteria of volatility and sign change frequency....232

Table 8.27 LSE: Spearman rank correlation analysis between forecast accuracy measures and indicators related to economic results....233

Table 8.28 Summary of forecast results from the ASE and the LSE series....235

LIST OF FIGURES

Figure 4.1 a,b Index price and Index returns plots of the ASE and the LSE series......71

Figure 4.2 a,b Frequency histograms of the ASE and the LSE series......72

Figure 4.3 a-d : Autocorrelograms and partial autocorrelograms for the ASE and LSE returns......73

e-f : Power spectrum of the ASE and the LSE series......73

Figure 4.4 a : Log(R/S) vs. Log(n) plot for the H exponent estimation of the raw and filtered ASE returns, the scrambled filtered returns and a Gaussian surrogate......94

b : V-statistic vs. Log(n) plot for the cycle-length estimation of the raw and filtered ASE returns, the scrambled filtered returns and a Gaussian surrogate......94

Figure 4.5 a : Log(R/S) vs. Log(n) plot for the H exponent estimation of the raw and filtered ASE 5-day returns......98

b : V-Statistic vs. Log(n) plot for the cycle-length estimation of the raw and filtered ASE 5-day returns......98

Figure 4.6 a : Log(R/S) vs. Log(n) plot for the H exponent estimation of the raw and filtered ASE 20-day returns......99

b : V-Statistic vs. Log(n) plot for the cycle-length estimation of the raw and filtered ASE 20-day returns......99

Figure 4.7 a : Log(R/S) vs. Log(n) plot for the H exponent estimation of the AR(1) filtered daily LSE returns and a Gaussian random surrogate 103

b : V-Statistic vs. Log(n) plot for the cycle-length estimation of the AR(1) filtered daily LSE returns and a Gaussian random surrogate......103

Figure 5.1a-h : Time series plots and Phase Space plots of the ASE daily returns, the LSE daily returns, a Gaussian random noise data and Lorenz attractor......108

Figure 5.2 Poincare map of a typical trajectory $\varphi_t(x)$ intersecting a cross-section S 109

Figure 5.3 a-b : Return map of a Gaussian random series 111

c-d : Return map of Lorenz attractor 111

Figure 5.4 a-b : Return map of ASE daily returns 112

c-d : Return map of LSE daily returns 112

Figure 5.5 ASE returns: Plot of the correlation dimension estimate (d) versus different embedding dimensions (m = 1,..,10) 114

Figure 5.6 LSE returns: Plot of the correlation dimension estimate (d) versus different embedding dimensions (m = 1,..,10) 121

Figure 5.7. ASE returns : convergence of the Largest Lyapunov Exponent for varying evolution time t_s = 5,10,15, 20 125

Figure 6.1. SVD analysis - Plot of the eigenvalues versus embedding dimension for the ASE returns 131

Figure 6.2. SVD analysis, ASE returns-Plot of the eigenvalues, with 95% confidence limits 132

Figure 6.3 a-j : The first 10 (out of 15) eigenvector plots for the ASE returns, after diagonalizing a (15x15) covariance matrix 133

Figure 6.4 a-h : 8 (out of 8) eigenvector plots for the ASE returns, after diagonalizing a (8x8) covariance matrix 135

Figure 6.5 a : ASE returns : Two dimensional phase-space reconstruction by the use of a (8x8) covariance matrix 136

b : ASE returns : Three dimensional phase-space reconstruction by the use of a (8x8) covariance matrix 136

Figure 6.6 a-c : SVD analysis - Plot of the eigenvalues versus embedding dimension for the noise filtered ASE series 138

Figure 6.7. SVD analysis, ASE returns-Plot of the eigenvalues for the noise filtered ASE series, with 95% confidence limits 139

Figure 6.8 SVD analysis - Plot of the eigenvalues versus embedding dimension for the LSE returns 140

Figure 6.9. SVD analysis, LSE returns - Plot of the eigenvalues with 95% confidence limits 141

Figure 6.10 a : LSE returns : Two dimensional phase-space reconstruction by the use of a (11x11) covariance matrix.......................................142

b : LSE returns : Three dimensional phase-space reconstruction by the use of a (11x11) covariance matrix.......................................142

Figure 6.11 SVD analysis - Plot of the eigenvalues versus embedding dimension for the noise filtered LSE series.......................................143

Figure 6.12 SVD analysis - Plot of the eigenvalues for the noise filtered LSE series, with 95% confidence limits 144

Figure 6.13 a-d : Autocorrelogram of the ASE and LSE noise filtered series, a Fractional noise series and Lorenz attractor.......................................146

e-h : Power Spectrum of the same series.......................................146

Figure 6.14 Plot of the correlation dimension estimate (d) vs. embedding dimension (m) for the noise filtered ASE series and its randomized phase surrogate, the noise filtered LSE series and the "short" noise filtered series.......................................149

Figure 6.15a: V-statistic plot of the ASE noise filtered series.......................................149

b: V-statistic plot of the LSE noise filtered series.......................................149

Figure 7.1. Graphical representation of the simplex method.......................................160

Figure 7.2. DVS plot of the logistic map for m=2,...10.......................................163

Figure 7.3 DVS plot of the ASE returns for m=2,...10.......................................163

Figure 7.4 DVS plot of the ASE filtered returns for m=2,...,10.......................................164

Figure 7.5 DVS plot of the ASE randomized sign returns for m=2,...10.......................................165

Figure 7.6 DVS plot of the Gaussian random surrogate of the ASE returns for m=2,...10.......................................165

Figure 7.7. Alternative version of the DVS plot, using correlation coefficient between actual and predicted values instead of the NRMS error.......................................166

Figure 7.8 a-c : Detection of chaotic behaviour through the investigation of the relationship between correlation coefficient and prediction time, for different series.......................................169

Figure 7.9 a-c : Log-log plot of (1-r) vs. T for different series.......................................172

Figure 7.10 a-e : Forecast NRMSE vs. embedding dimension for different series, as a test for an upper dimensionality bound 174

f-j : Correlation coefficient between actual & predicted values vs. embedding dimension for different series as a test for an upper dimensionality bound..174

Figure 7.11 a-e : Average forecast NRMS error vs. embedding dimension for different series as a test for an upper dimensionality bound...................175

Figure 8.1 a-h : Forecast results comparisons of each different model for alternative scenarios based on different library and prediction period length...202

Figure 8.2 : ASE: Time series plot of the 1987 annual sub-period.............................214

Figure 8.3 : LSE: Time series plot of the 1987 annual sub-period.............................231

SYNOPSIS

This dissertation presents an effort to implement nonlinear dynamic tools adapted from chaos theory in financial applications. Chaos theory might be useful in explaining the dynamics of financial markets, since chaotic models are capable of exhibiting behaviour similar to that observed in empirical financial data.
In this context, the scope of this research is to provide an insight into the role that nonlinearities and, in particular, chaos theory may play in explaining the dynamics of financial markets.

From a theoretical point of view, the basic features of chaos theory, as well as, the rationales for bringing chaos theory to the attention of financial researchers are discussed. Empirically, the fundamental issue of determining whether chaos can be observed in financial time series is addressed.
Regarding the latter, empirical literature has been controversial. A quite exhaustive analysis of the existing literature is provided, revealing the inadequacies in terms of methodology and the testing framework adopted, so far.

A new "multiple testing" methodology is developed combining methods and techniques from the fields of both Natural Sciences and the Economics, most of which have not been applied to financial data before. A serious effort has been made to fill, as much as possible, the gap which results from the lack of a proper statistical framework for the chaotic methods. To achieve this the bootstrap methodology is adopted. The empirical part of this work focuses on the comparison of two markets with different levels of maturity; the Athens Stock Exchange (ASE), an emerging market, and London Stock Exchange (LSE). Our aim is to determine whether structural differences exist in these markets in terms of chaotic dynamics.

In the empirical level we find nonlinearities in both markets by the use of the BDS test. R/S analysis reveals fractality and long term memory for the ASE series only. Chaotic methods, such as the correlation dimension (and related methods and techniques) and the largest Lyapunov exponent estimation, cannot rule out a chaotic explanation for the ASE market, but no such indication could be found for the LSE market. Noise filtering by the SVD method does not alter these findings. Alternative techniques based on nonlinear nearest neighbour forecasting methods, such as the "piecewise polynomial approximation" and the "simplex" methods, support our aforementioned conclusion concerning the ASE series.

In all, our results suggest that, although nonlinearities are present, chaos is not a widespread phenomenon in financial markets and it is more likely to exist in less developed markets such as the ASE. Even then, chaos is strongly mixed with noise and the existence of low-dimensional chaos is highly unlikely. Finally, short-term forecasts trying to exploit the dependencies found in both markets seem to be of no economic importance after accounting for transaction costs, a result which supports further our conclusions about the limited scope and practical implications of chaos in Finance.

Introduction

The mechanisms explaining the dynamics of asset prices and returns is a fundamental issue in finance. The traditional approach, which assumes that fluctuations in these variables are largely the result of random processes that can be represented by linear stochastic models, has been empirically disputed and there has been a growing interest in nonlinear approaches. A major part of this literature is dealing with various types of nonlinear stochastic specifications among which we find the very popular ARCH-type models. Recently, chaos theory, which describes a type of (deterministic) nonlinear behaviour, has attracted the attention mainly because of the ability of chaotic models to generate time paths similar to those observed in empirical financial data.

The question of whether chaos theory may be proven to be an important mathematical tool in getting a better understanding of the financial markets is still unresolved. Many authors have employed methods and techniques from the nonlinear dynamics field to analyse different economic and financial series and try to determine whether chaotic behaviour is present. Up to now, most empirical findings have been contradictory. What is not clear is whether the source of this controversy is due to: (i) the nature of the different data sets employed (i.e. sample size, noise level, aggregation method, prefiltering processes etc.), (ii) the different methods and techniques that have been used, (iii) the improper use of the chaotic testing framework, (iv) the fact that most chaotic techniques are not statistical in nature which makes the interpretation of the results more difficult (even arbitrary) or (v) combinations of these sources.

In this context, the main purpose of this doctoral dissertation is to investigate the possibility of a deterministic nonlinear structure in financial data, thus providing insights towards the role that chaos theory might play in explaining financial markets' dynamics. Within this framework an extensive empirical analysis is pursued, implementing nonlinear dynamic tools adapted from chaos theory in financial series that have not been examined before.

The original contribution of this work is two-fold in scope:

From a methodological perspective, the innovation lies in introducing a "multiple testing" methodology which combines the disciplined approach of the Natural Sciences - in the application level - with a variety of different methods and techniques available in the Economics and the Natural Sciences fields, some of which have not been used before with financial data. In this respect we introduce: the Singular Value Decomposition (SVD) method for phase space reconstruction and noise filtering purposes, "Theiler's" specification, "phase randnomisation" and "sign randomisation" techniques in the context of the correlation dimension estimation, the "independent realisations" method, as well as, the "DVS plot", the "varying prediction time" and the "dimensionality" techniques that can distinguish between random and chaotic specifications, based on nonlinear forecasts. The nonlinear methods adopted are also compared to various linear specifications in terms of their ability to produce economic results in a simulated trading environment. Finally, a serious effort has been made to enhance the robustness of some of the methods employed by using the bootstrap method for statistical inference purposes.

From an empirical application standpoint the above principals and ideas are applied to a mature market (London stock exchange) and for the first time to an emerging market (Athens stock exchange). The motivation for this comparative study lies with the fact that emerging capital markets (ECM) are very likely to exhibit characteristics different from those observed in developed capital markets. The first have received little attention so far in the literature and are often less efficient and probably more suitable for chaotic analysis since the underlying moving forces might be fewer than the ones in the mature and more efficient markets.

This dissertation is organised as follows:

Chapter 1 provides an introduction to the general theoretical framework of chaos theory and nonlinear dynamics. New concepts and notions related to this framework, such as chaotic maps, attractors, fractals, dimensions and Lyapunov exponents, are presented and discussed. The approach followed to present the above-mentioned issues is both technical and intuitive in order to give a brief but also comprehensive picture of the theory behind the empirical framework of this research. In the same Chapter the alternative methods of reconstructing a phase space of a dynamical system from a single variable are presented. This is one of the most important theoretical and empirical issues in the analysis of dynamical systems and provides a basic part of the empirical analysis of this study. Specifically, the "method of delays" and the "Singular Value Decomposition method" are analytically discussed, the latter being also a valuable tool for noise reduction purposes.

Chapter 2 includes the chaotic testing framework of this research. The basic methods and techniques to be employed in the empirical part of the study are analytically presented. For each method, its origins, technical description, advantages and shortcomings as well as application details are discussed.

Chapter 3 motivates the use of the nonlinear dynamics and chaotic approach in the markets. A theoretical perspective is provided, covering and comparing the traditional stochastic framework with the new chaotic alternative. The rationales for bringing chaos theory to the attention of financial researchers are discussed and evidence from the relevant empirical literature is presented. Finally, data validation issues and the adopted methodology for the empirical part of this research are presented, too.

In Chapter 4, the empirical applications of this study start with the statistical description of the two data sets and tests for nonlinearity and fractality, which in our methodology are the preliminary steps of the chaotic analysis to follow.

Empirical evidence from the application of the chaotic methods in our data is presented in Chapter 5.

In Chapter 6, the empirical findings of the previous Chapter are cross-examined via the application of the SVD method. The methodology adopted here is used with financial data for the fist time and it is shown that it can be a valuable tool in the nonlinear dynamics empirical framework.

Chapters 7 and 8 are devoted to forecasting methods and applications. In Chapter 7, nonlinear forecasting techniques have been used as additional diagnostic tools for distinguishing between chaotic and stochastic specifications.

Chapter 8 presents our efforts for short-term prediction in both markets examined, by the use of alternative linear and nonlinear models. The most interesting part is the economic assessment of our forecasts through trading strategy simulations that permit to draw very useful conclusions about the ability to make profits in "almost" real market conditions and to assess the performance of techniques that have been developed in the nonlinear dynamics framework.

Finally, a summary of the findings and the final conclusions of this research are presented in Chapter 9.

Chapter 1

THE THEORETICAL FRAMEWORK OF CHAOS THEORY AND NONLINEAR DYNAMICS

1.1. INTRODUCTION

In the last few years, the exploding developments in informatics and computing ability have made possible the in depth exploration of complex systems mainly in the Natural Sciences. A central issue in this research has become the unravelling of the structure of "disorder", a common and usually prevailing, yet neglected characteristic of many dynamical systems. The exploration of nonlinear systems aiming at a new understanding of the "laws of disorder" has been dubbed "the Chaos theory" or "Nonlinear Dynamics". In fact, as we will see, Chaos is a subset of the more general nonlinear dynamics framework.

Chaos can be considered as the latest scientific revolution or using Kuhn's (1962) terminology as a “paradigm shift” from the still prevailing in most scientific fields “linear paradigm” to the “nonlinear paradigm”. It is a "new science" [Gleick (1987)] which, for some of its advocates, is equally revolutionary to relativity and quantum mechanics. According to some authors [Prigogine (1980), Prigogine and Stengers (1984), Cvitanovic (1984), Briggs & Peat (1989)], Chaos can be considered as a generalised paradigm in Nature.

Chaos theory has been developed in the theoretical framework of Dynamical Systems Analysis and draws heavily on specialised disciplines like Differential Topology, Fractal Geometry etc. It cuts across the interdisciplinary dichotomy between "determinism" and "randomness", which characterises the way that traditional science approaches the various phenomena. Actually, these two apparently opposite modes of behaviour seem to reconcile in the Chaos theory framework.

The development of Chaos theory is an interdisciplinary effort. So, it is not surprising that it has been proven to be very useful in analysing the behaviour of many phenomena, the study of which has been the subject of a wide range of different disciplines like Physics, Biology, Meteorology, Chemistry, Medicine, Ecology and Economics and Finance.

The early ideas of the scientific approach to Chaos date back to the last Century. The famous physicist Ludwig Boltzman in 1870 was the first to prove that the observed Chaos in Thermal Entropy was just another expression of the Newtonian order [Bai-Lin (1984)].

A little later, Henri Poincare discovered that even a simple non-linear system has an inherent potential to create Chaotic behaviour characterized by what we call today SDIC (Sensitive Dependence on Initial Conditions).

The modern study of Chaos dates back to the 1960s. In a brief and arbitrary reference to the major steps of this research we should mention:

Lorenz (1963), a meteorologist who re-discovered the SDIC principal (commonly known as "The Butterfly Effect") in a simple non-linear dynamical model for weather prediction consisting of three differential equations that produce the famous “Lorenz attractor”.

Smale (1963), a mathematician specialised in Topology who was the first to connect Topology with Chaos theory. His innovation was a topological transformation known as Smale's "horseshoe", which provides a base for understanding the chaotic properties of dynamical systems.

York [Li and York (1975)] and May (1976), a mathematician and a biologist respectively who studied a very simple but now the most famous chaotic system known as the "Logistic equation" or the "Logistic map". This simple system has been extensively used in the literature for educational [Savit (1988), Ott (1981), Ott et.al. (1990)] and modelling [Malliaris and Philippatos (1992)] purposes. It should be also mentioned that York was the one who gave to Chaos its very name.

Feigenbaum (1978,1979), who discovered Universal Laws hidden in the routes to Chaos. Feigenbaum proved that there are structures in non-linear systems that are always the same, which means that chaotic routes obey to the same Universal Laws.

To close this short chain of the early discoveries related to Chaos, we should mention Mandelbrot (1982), who developed the notion of Fractals or self-similarity across scales. This concept plays an important role to the description of "attractors", the basic dynamical object under consideration in the Chaotic framework.

1.2. DYNAMICAL SYSTEMS AND CHAOS DEFINITION

Chaos is closely related to the analysis of Nonlinear Dynamical Systems. Dynamical systems are defined formally in terms of a function relating the system's past to its future and vice versa. This function expresses the state transition rule, that is, the rule the system uses to make transitions from one state to the next. Since this rule is a mathematical function a dynamical system is a deterministic one. Moreover, this function is generally nonlinear, thus nonlinear deterministic dynamical systems are of our interest.

The dynamical equations or equations of motion of a dynamical system in time t, can be of the form :

$$dx/dt = f(x) \qquad (t \in R^1) \qquad (1.1) \text{ or}$$

$$x \rightarrow g(x) \text{ or } x_{t+1}=g(x_t) \qquad (t \in Z) \qquad (1.2)$$

where $x \in R^n$ the n-dimensional Euclidean space. Each $x =(x_1,x_2...)$ represents a state of the system, i.e. a point in a space of n-dimensions R^n, which is called the ***phase (or state) space*** of the system and can be defined as the one (or multi)-dimensional space where the dynamic evolution of a system can be represented and observed.

We refer to (1.1) as a vector field or ordinary differential equation (o.d.e.) and to (1.2) as a map or difference equation. In (1.1) the time t is a continuous variable while in (1.2) the time variable t, t=0,1,2,..., is a discrete one.
The solution of a dynamical system is called the ***orbit or the trajectory*** of the system. Orbits produced by o.d.e. systems are continuous curves, while orbits of maps are discrete sets of points.

A solution curve of a continuous-time dynamical system is denoted by the ***flow*** $\varphi(t, x_0)$ which provides the value of x at a time t given an initial condition x_0, i.e. $\varphi : x_0, t \rightarrow x_t$, where x_t is the new position of the initial state of the system after time t.

If $M \subset R^n$ is a set invariant[1] under the flow $\varphi(t,x)$ generated by (1.1), i.e. $\varphi(t,M) \subset M$ for all $t \in R$, then M is said to be ***chaotic*** if:

1. The flow $\varphi(t,x)$ has ***sensitive dependence on initial conditions*** on M. Formally, if there is a distance $\delta>0$ arbitrarily small such that for any $x \in M$ and any neighbourhood[2] U of x, there exists $y \in U$ and $t>0$ such that $|\varphi(t,x)- \varphi(t,y)|>\delta$. In other words for any point $x \in M$ there is (at least) one point arbitrarily close to M that diverges from x. This rate of divergence for many authors is reqired to be exponential, however, according to Wiggins (1990) there are cases of experimental chaotic systems exhibiting nonexponential contraction or expansion rates.

2. The periodic orbits of $\varphi(t,x)$ are ***dense*** in M. The property of density entails that for any initial condition x(0) there exists another x(0)′ arbitrarily close to the former, that is periodic. This means that in chaotic maps regularity (periodicity) and structure co-exist with unpredictability.

3. The flow $\varphi(t,x)$ [resp. the map g(x)] is ***topological transitive*** on M. Formally, topological transitivity can be defined as :

For any two open sets $U,V \subset M$, $\exists\, t \in R \ni \varphi(t,U) \cap V \neq \varnothing$

Roughly speaking topological transitivity means that there is a trajectory which visits any arbitrarily small preassigned region in the set. This property guarantees non-decomposability of the set i.e. non-separation into subsystems that behave independently.

The above is a mathematical/topological approach in defining chaos, given by several authors in the literature [e.g. Devaney(1989,1990), Winnie(1992)].

---

1 A set M is said to be invariant under a vector field if for any $x_0 \in M$ we have $x(t,t_0=0,x_0) \in M$ for all $t \in R$. A differentiable invariant set can have the structure of a differentiable manifold which is a set having locally the structure of Euclidean space. In linear settings a manifold is a linear vector subspace of R^n. In nonlinear settings a manifold is a d-dimensional surface M placed in an n-dimensional Euclidean space R^n, defined by the relations between the coordinates of R^n i.e it can be locally represented as a graph.

2 The neighbourhood of a point x, is an arbitrarily small open set containing x.

However, a rigorous and all encompassing definition of Chaos is still lacking. Several non-technical definitions can be also found in the literature, defining Chaos in a rather general way. Among them Larain (1991a) defines Chaos as an "irregular complex behaviour that seems random but actually has some hidden order", Theiler (1990a) and Ramsey, Sayers and Rothman (1990) define it as the irregular behaviour generated by simple nonlinear functions and Shaffer (1991), Hsieh (1991) and Peters (1991a) connect chaos to the notion of determinism and define it as as a non-linear deterministic process that can produce random looking results.

A similar definition to the latter has been suggested for wider acceptance in a conference on Chaos [Royal Society, London (1986)]. According to this, "Chaos is the stochastic behaviour occurring in a deterministic system", where a deterministic system is considered to be one which comprises no exogenous random variables
Finally, some authors [De Grauwe and Vansanten (1990), Savit (1989), Barnett and Chen (1988) and Ruelle (1990)] define Chaos as the prevelence of SDIC, considered to be its most important characteristic property.

A working definition of Chaos can be applied to the analysis of a time series i.e. the observable output of a system that we can measure (scalar quantity). The essential requirements for a chaotic series are that the underlying behaviour is generated by a discrete or continuous time nonlinear model which is deterministic and that the time path of these series will exhibit SDIC.

According to Brock and Sayers (1988) the series $\{a_t\}$ has a deterministic chaotic explanation if there exists a system (h,F,x_0), such that h maps R^n to R (h is the measurement function), F maps R^n to R^n, $a_t=h(x_t)$, $x_{t+1}=F(x_t)$ and x_0 is the initial condition at t=0. The map F is deterministic, the phase-space is n-dimensional and all trajectories $\{x_t\}$ generated by the system lie on a subset A of the phase space called the attractor of the system, and two nearby trajectories on A locally diverge exponentially.

There are invariant quantities[3] (measures), such as Dimensions and Lyapunov exponents, that can quantify the properties of an attractor.

These measures can be calculated from a series and provide ways to characterize the nature of the underlying system. For example, a positive largest Lyapunov exponent is used [Bask (1998)] as an operational definition of deterministic chaos if the dynamical system generating the time series is dissipative. All these concepts related to the practical objective of identifying chatic processes will be further analyzed in what follows.

---

3 Invariance is related to the property of these quantities to assume the same value irrespectively of the measurement procedure, the coordinates chosen etc. Dimensions are also referred in the literature as "static invariants" that take different values for the deterministic and non-deterministic case respectively, enabling to distinguish between the two cases. On the other hand Lyapunov exponents are reffered to as "dynamic invariants" related to the exact nature of the attractors.

1.3. ATTRACTORS

Given the dynamical systems described in the previous section by the equations (1.1) and (1.2), a close invariant set $A \subset R^n$ is called an ***attracting set*** if there is some neighbourhood U of A such that : $\forall\ x \in U$, $\forall t \geq 0$, $\varphi(t,x) \in U$ and $\varphi(t,x) \rightarrow A$ as $t \rightarrow \infty$

The closure of the set of all initial conditions satisfying the above definition is the ***domain or basin of attraction*** of A and is given by : $\bigcup_{t \leq 0} \varphi(t,U)$

An ***attractor*** is a topologically transitive attracting set A that according to the above definition has a neighbourhood U such that the flow $\varphi(t,x)$ of the system comes arbitrarily close to all members of A as the system evolves. If A is chaotic (see definition in the previous section) it is called a ***strange attractor*** [Eckman(1981), Wiggins(1990)].

Attractors are related to ***dissipative dynamical systems*** as opposed to ***conservative*** ones. According to Shaw (1981), from the "Information Theory" point of view, dissipation implies the loss of memory of initial conditions once the system has reached its asymptotic regime. In dissipative systems what matters is the long-term behaviour of the system (or its asymptotic regime), which is the only observable behaviour since it is repeated. Hence, we can say that dissipation permits to observe the permanent behaviour of the system and to ignore its transient behaviour.

In several non-technical definitions in the literature an attractor is defined as a set on which experimental points (generated by a dissipative dynamical system) accumulate for large t [Ruelle (1983)], as a subset of an n-dimensional phase space towards which almost all sufficiently close trajectories get attracted asymptotically [Grassberger and Procaccia (1983a) and Brock and Sayers (1988)], as the low dimensional set onto which the flow of a dissipative system contracts as the system evolves [Broomhead and King (1986b)], as the subset of points towards which any dynamical path will converge [Ramsey et.al. (1990)] or as the equilibrium level of a system, i.e. the level that the system reverts to, after the effects of perturbation of the system die-away Peters (1991a).

Attractors might be of several types. Simple examples of (nonchaotic) attractors are fixed points (after the transient time the system settles to a stationary state) and limit cycles(the system approaches a periodic motion). A strange attractor is the most interesting type from our point of view since it characterises chaotic systems. Strange behaviour in time manifested in the SDIC property of an attractor has its counterpart in its geometry, which is very complicated. So strangeness is also related to the fractal nature (or fractal geometry) of an attractor [Grebogi et. al. (1984)].
In topological terms, the fractal nature of a strange attractor can be described as an infinitely folded sheet of infinite extend located in a bounded region. In other words a strange attractor arises according to a stretching and folding, at the same time, procedure. This process is the underlying mechanism producing both the fractal structure and the chaotic conditions (SDIC) on the attractor and results to an unstable final motion generated by the system's dynamics within the attractor.

Fractals are objects in which the parts are in some way related to the whole and its basic characteristic is "self-similarity" or "symmetry across scales" [Mandelbrot (1982)]. Fractals can be either deterministic (e.g. fractal shapes or mathematical fractals like the Candor set) or random (e.g. natural objects, time series etc.). Random fractals are generated by combinations of deterministic rules chosen at random at different scales thus combining "randomness" with "determinism". They usually have parts, which are qualitatively related in the sense that they might not look self-similar, but have similar statistical characteristics at different scales. This latter feature of fractals is called “scale invariance”. Characteristic scaling described by a power law and invariace across scales are the two basic characteristics of fractals that can be identified in an empirical context. Most of the chaotic systems that have been investigated are random fractals. Typical chaotic systems have strange attractors with fractional (noninteger) dimensions. Dimensions provide ways to quantify self-similarity and increase our knowledge about a system as will be shown in the section to follow.

1.4. Dimensions

Dimensionality in nonlinear dynamics is related to information that can help to describe a dynamical system, chaotic or not. In this context dimensions can provide information about the number of the dynamic variables governing the motion of a system (or in empirical terms about the number of variables needed to model the system). In topological terms, dimensions provide information related to the folding procedure needed to generate an attractor.

In empirical terms, dimensions are complexity measures which can specify the effective degrees of freedom of a dissipative dynamical system by measuring the dimension of its attractor. A strange attractor which characterises a chaotic system should exhibit low-dimensionality (a few active modes), while a stochastic system should have many active modes, i.e. it is a high-dimensional one. Hence, this dimensionality difference could help distinguishing between these two modes of behaviour provided that we can measure the dimension.

There are several different notions of dimensions, conceptually and mathematically interrelated, which are studied in the context of Dimension Theory, a quantitative field that draws heavily on differential topology, theory of sets and differential calculus.

Dimension measures to be presented here are **Fractal (or Hausdorff-Besikovich) dimension, and Correlation Dimension,** the first because it is very useful in undertanding this new dimensionality consept and the second because it is the most impotrant from an empirical point of view.

1.4.1. The Fractal Dimension

Fractal Dimension (FD) is a subversive concept to the Euclidean geometry. Fractal objects (or shapes) are often rough, discontinuous, irregular and mathematically non-differentiable. Consequently, Euclidean (or topological) dimensions cannot be used to describe them and another measure of "complexity" is needed. This measure should be capable of measuring difficult to define qualities in an object, like irregularity or roughness. FD is such a measure, its basic characteristic being that it may not take integer values[4].

Intuitively, FD can be defined as the degree of irregularity corresponding to the efficiency of an object to fill its space and is the product of all the factors influencing the system (its degrees of freedom) that produce the object. A non-integer FD indicates that the orbits of a system tend to fill up less than an integer subspace of its phase space. The association between FD and chaotic systems stems from the basic stretching and folding mechanism producing the chaotic behaviour. It has been shown [Shaw (1981)] that chaotic attractors of three-dimensional continuous time dissipative systems have a FD between 2 and 3.

FD can be calculated by counting the number of circles of radius r needed to cover the fractal shape. As the radius r is increased, the number of circles scales exponentially to the radius, according to the following relationship :

$$N(2r)^D = 1 \qquad (1.3)$$

from which FD is derived as:

$$D = \log N/\log(1/2r), \qquad (1.4)$$

where: N = number of circles, r = radius and D = Fractal Dimension.

Following this formula, fractal shapes like the "Koch snowflake" [Mandelbrot (1982)] has a Fractal Dimension D= 1.26, and the Cantor set has D= 0.631.

This calculation can be generalised for the case of a multi-dimensional fractal object (or an attractor) by replacing the circles with hyper-cubes or hyper-spheres, the dimensions of which depend upon the dimensions of the object (or the attractor).

Following a more formal approach used in Schuster (1988) and in Eckmann & Ruelle (1985) we can consider a set of points S, where S is a compact metric space[5] and

4 FD is an integer for "nice" cases (e.g. clear Euclidean lines and shapes), and a non-integer for "irregular" ones.

5 A metric space is a set for which a measure of distance between points and the neighbourhood of a point can be defined.

$N(r,S)$ the minimum number of open balls (or spheres)[6] of radius r needed to cover S, then:

As $r \to 0$, $$N(r,S) = Ar^{-D} \quad (1.5),$$

where A is a constant and the quantity D can be defined as:

$$D = \lim_{r\to 0}[\log N(r,S)/\log(1/r)] \quad (1.6)$$

while away from the limit (1.6) becomes:

$$\log N(r,S) \overset{r\to 0}{=} D\log(1/r) + \log A \quad (1.7)$$

The quantity D is called the (Kolmogorov) **Capacity Dimension** (or simply the Capacity, or the Limit Capacity) of S.

The capacity measure D is theoretically equal to or smaller than the Fractal or ***Hausdorff dimension*** D_H ($D_H \geq D$), but practically they are very close to each other and empirically it is difficult (and rather meaningless) to be distinguished.

Capacity and Hausdorff dimensions are very useful measures of the local structure of attractors. Yet, they suffer from certain disadvantages.

The first disadvantage is related to the fact that they are both geometric measures, that is, they do not take into account the frequency with which the various parts of an attractor are visited by a typical trajectory of the underlying dynamical system. Hence, they are metrics and not probabilistic measures, which are more appropriate to describe a strange attractor.

The second disadvantage is the need of a box-counting algorithm in order to calculate them, a method highly impractical and costly in empirical applications, especially when the system does not have a very low dimension.

1.4.2. The Correlation Dimension.

From an empirical point of view, the Correlation Dimension is the most important dimension measure, due to its computational convenience. The underlying idea to the Correlation Dimension is related to the spatial correlations between the random points

---

6 If a plane is covered instead of a multi-dimensional object, only a two-dimensional slice (i.e. a circle) of the sphere is used. In the case of an object or an F-dimensional attractor, F-dimensional hyper cubes (instead of spheres) of side length r can also be used to define the fractal dimension.

on an attractor. Specifically, if we consider a set $\{X_i,\ i=1,\ldots N\}$ of points on a strange attractor, most pairs of (X_i,X_j) points with i=j will be dynamically uncorrelated due to the SDIC property of the attractor (i.e. the exponential divergence of the trajectories). Nevertheless, all points are bounded to be on the attractor, which means that they will be temporally uncorrelated but spatially correlated. Based on this idea, Grassberger and Procaccia (1983a) (hereafter G-P) suggested an alternative way to estimate the N(r,S) cover in (1.5) by replacing the "box-counting" algorithm with the ***G-P Algorithm*** or ***Correlation Integral*** or ***Correlation Sum***, defined as:

$$C(r) = \lim_{N\to\infty} 1/N^2 \ \{\text{number of pairs (i,j) whose distance } |X_i - X_j| \text{ is less than } r\} \qquad (1.8)$$

where N is the number of points on the attractor and the distance is usually estimated by the use of a suitable distance norm.

G-P have established that the Correlation sum for very small r (i.e. $r\to 0$) scales like a power law:

$$C(r) = r^d \qquad (1.9)$$

This correlation exponent d is a useful measure of the local structure of a strange attractor and a tool to quantify self-similarity.

Taking the logarithms, (1.9) can be also written as:

$$d = \lim_{r\to 0}[\log C(r)/\log(r)] \qquad (1.10)$$

where d is called the ***Correlation Dimension*** or ***the Correlation Exponent***.

Correlation dimension is a probabilistic measure since it is by construction sensitive to the frequency by which the regions of the attractor are visited. This property together with its less costly way of calculation renders it the most attractive dimension measure in empirical applications. The procedure that has to be followed in order to define the correlation dimension of an attractor suggests the possibility of using the correlation dimension estimates in order to distinguish between random and deterministic systems. A detailed picture of the application issues with respect to the correlation dimension estimation will be given in the next Chapter.

1.5. Lyapunov Exponents

Lyapunov Exponents or ***Characteristic Exponents*** (hereafter LE) form another category of measures which, together with dimension measures, can be considered as providing the most useful information in order to describe a non-linear dynamical system.

LE are measures of the SDIC property of chaotic systems and measure the average rate at which nearby trajectories of a dynamical system diverge or converge over time in phase space. In topological terms and in the case of chaotic systems, LE describe the stretching (or divergence) in phase space needed to generate a strange attractor.

Each dynamical system has a spectrum of LE, which might contain positive, negative and zero exponents. This spectrum corresponds to the number of dimensions in phase space of the system (each dimension has its own Lyapunov exponent).

LE offer a way to classify attractors and this ability is related to their signs, which provide a qualitative picture of a system's dynamics. In general, negative exponents correspond to "contraction"[7], zero exponents correspond to convergence or divergence to a slower rate than exponential, and positive exponents correspond to expansion or divergence at an exponential rate.

A dynamical system with an attractor should have at least one negative exponent, since contraction in some directions is necessary in order to have the attractor generated, and the sum of the LE is also negative[8].

A system with only negative LE contracts to a fixed point (i.e. it has a point attractor). In any other case a system will have at least one zero valued Lyapunov Exponent and if it has one or more positive LE it is characterised as a Chaotic one (i.e. its attractor is a strange one).

---

7 How long it takes for a system to return to its attractor after it has been perturbed.

8 In dissipative systems, the volume occupied by the attractor with respect to the volume of phase space is in general very small (practically an attractor is a zero volume limit set). So we can say that a dissipative system contracts the phase space volumes (they are contracted by the time evolution) but this contraction may not happen in all directions. Some of them might be stretched and this stretching property corresponds to chaotic behaviour i.e. to exponentially diverging trajectories and to positive LE. Yet the non-stretched directions are so much contracted that the final volume of the attractor becomes smaller which means that contraction prevails and this is reflected to the negative sum of the LE.

Using more technical terms, there are different ways to define LE. According to Ruelle (1984), for a dynamical system f(x(t)) SDIC is present if :

$$|\delta x(t)| - |\delta x(0)| = e^{L(t)} \quad , \ L>0 \qquad (1.11)$$

where $\delta x(0)$ represents an infinitesimal change to system's initial conditions and $\delta x(t)$ represents the corresponding change at time t. Obviously, L gives the exponential rate of the change and is a Lyapunov exponent.

According to Eckman & Ruelle (1985), LE can be defined as the exponential rate at which a slight perturbation in one of the n directions of the time evolution of a dynamical system increases (or decreases) with time. This exponential growth can be locally measured by the eigenvalues of the Jacobean of the flow of the system. Formally, the growth of the perturbation δx about a point x is given by:

$$\delta x(t) = (D_x \varphi_t)\delta x(0) \qquad (1.12)$$

where $D_x \varphi_t$ is the Jacobean of the flow φ_t and $\delta x(0)$ and $\delta x(t)$ represent the (slight) change in initial conditions and the change in time t respectively.

The Lyapunov exponents L_i are given by the asymptotic growth rate of the eigenvalues of the Jacobean, as:

$$L_i = \lim_{t \to \infty} 1/t \ | \ D_x \varphi_t \ | \qquad (1.13)$$

A third approach which we adopt in our empirical applications is due to Wolf et al. (1985) who relate them to the growth rate of an infinitesimal volume element (i.e. an n-sphere of initial conditions for a system in an n-dimensional phase space).

As the system evolves under the flow, the n-sphere will be stretched to an n-ellipsoid since the nearby points in the sphere representing different initial conditions will diverge over time. The infinitesimal volume of the sphere permits a linearisation and, if we denote by p_i the length of principal axes of the ellipsoid (i.e. the i-th dimension of the system), we can define LE as:

$$L_i = \lim_{t \to \infty} (1/t) \log_2 [(p_i(t)/p_i(0)] \qquad (1.14)$$

where L_i represents the i-th Lyapunov Exponent for the i-th dimension (p_i) and L_i are ordered from the largest to the smallest.

LE, though very important in describing a dynamical system, can be accurately calculated only when the equations of motion are known and, even then, it is difficult to calculate all of them. Yet, a very useful theorem by Oseledec (1968), called the

Multiplicative Ergodic Theorem, proven also by Johnson, Palmer and Sell (1984), states that (almost) all LE have the Largest one as a limit. The empirical implications of Oseledec's theorem are of major importance since the calculation of the LLE only suffices to characterise an attractor. That is, a negative, zero or positive LLE of a dynamical system renders its attractor as a point one, a periodic orbit or a strange (chaotic) attractor respectively.

Another aspect of interest is the quantitative information provided by the magnitude of the LE, further to the qualitative information provided by their signs. This quantitative information is given in Information Theory terms developed by Shannon [Shannon and Weaver (1963)]. According to this approach, Lyapunov exponents measure the rate at which the system creates or destroys information, as it evolves. That is, LE are closely related to the ***Entropy*** of the system. In general, a chaotic system generates new information which increases the entropy of the system and a certain relationship associates an Entropy measure, called the Kolmogorov entropy to LE:

$$K \leq \Sigma_i L_i^+ \qquad (1.15)$$

Kolmogorov entropy (K_2) [Kolmogorov (1961)] measures an aggregate of the stretching behaviour of a system while LE quantify the stretching and contracting in various directions. K is zero for a quasi-periodic system, infinite for a random iid system and finite for a chaotic one.

Normally, LE are expressed as "bits" of information[9] per orbit or per iteration in the case of continuous and discrete time systems, respectively. Empirically, this can be translated into a loss of predictive ability as to where the system will be after certain orbits or iterations. For example, in the case of a well-known attractor that has been studied analytically, namely the Henon attractor [Henon (1976)], the LLE is 0.42 bit. If our measurement accuracy is 2 bits, all information (and our predictive ability) will be lost after 2/0.42 = 4.8 iterations. After that point we shall not be able to specify the state of the system and what we know is just that it is somewhere on the attractor. We shall see in the following Chapters how this kind of approach can be useful in analysing financial data.

9 This is why in equation (1.14) $\log_2$ is used instead of $\log_e$.

1.6 PHASE SPACE RECONSTRUCTION METHODS

A first necessary step in the characterisation of an attractor by the use of certain metrics, such as Dimensions and Lyapunov Exponents, is the construction of the phase space of the system or, in terms of observability, the construction of the phase portrait of the trajectories generated by the system. In the case that the governing equations of the system (its law of motion) are a priori known, this step can be accomplished in a rather straightforward manner. This is the case for simple chaotic maps and many other cases of known attractors generated by specific and a priori known systems of o.d.e. (such as Rossler's attractor [Rossler (1976)], Lorenz's attractor [Lorenz (1963)] etc.) that have been in depth analysed in the Physics literature

Nevertheless, in practice we observe only a single degree of freedom (one dynamical variable), that is, a time series, which is assumed to be generated by the dynamical system. The issue of reconstructing a system's full motion in phase space from the observation of a single degree of freedom is the cornerstone of the empirical investigation of unknown dynamical systems. This is because a proper reconstruction is a sine-qua-non prerequisite for any analysis to follow.

The basic method for the phase space reconstruction, called the ***"Method of Delays"(MOD)***, has been developed separately, but almost simultaneously, by Takens (1980) and Packard et al. (1980) and has been employed by almost all empirical applications in the field.

An alternative method is the ***"Singular Value Decomposition" (SVD)*** method, developed by Broomhead & King (1986a). According to Kugiumtzis and Christophersen (1994), under certain conditions relevant to reconstruction parameters setting, the two methods perform equivalently. Yet, the same authors show that in case of short and noisy data, the SVD method outperforms MOD. Both these methods will be presented next.

1.6.1. The Method of Delays

Lets consider a dissipative dynamical system evolving in a n-dimensional phase space R^n and x_t ($x_t \in R^n$) the state vector of the system, which represents the state of the system at time t. The system evolves according to a law of motion which can be represented as:

$$F : R^n \rightarrow R^n, \quad \text{that is, } x_{t+1} = F(x_t) \text{ for all t} \qquad (1.16)$$

As the system evolves, the trajectories $\{x_t\}_{t=1..\infty}$ generated by the system will be attracted by a subset A of the state space (i.e. the attractor of the system) which in general will have dimension d, lower than that of the state space (i.e. d<n), where d corresponds to the active degrees of freedom of the system.

The problem we face in practice is that we know nothing about the real deterministic system (the law of motion), the state vector x and the dimensionality of the state space. Instead, we observe an experimental signal: $a_t = h(x_t)$, where a_t is a scalar variable observed during period t, i.e. measurements corresponding to a data path $\{a_t\}_{t=1..\infty}$, which is assumed to depend on the state vector. In other words the evolution of a single variable should be affected by all the degrees of freedom of the system.

This dependence is expressed through a function $h(x_t)$, called the ***"observer (or the measurement) function"***, which is also unknown to the observer and relates the unknown state vector x to the single observable a_t.

Takens (1980) has proved that there is an extractable relationship between the hidden dynamics of a system and the observed variable. In other words, although we know nothing about the real system, we can reconstruct its dynamics in order to investigate whether a deterministic explanation exists for the observed data path.

This is possible by employing the method of delays, which entails the generation of several different signals (series) from the original one. These signals are obtained by the use of ***time delays*** of the $\{a_t\}$ signals which are also called ***"past histories"*** or ***"m-histories"***, and can be written as:

$$\Phi_m(x_t) = \{h(x_t), h[F^{\tau}(x_t)], h[F^{2\tau}(x_t)], \ldots, h[F^{(m-1)\tau}(x_t)]\} \qquad (1.17)$$

where τ is an arbitrary constant called the ***delay time*** and m is called the ***embedding dimension***.

It is obvious that (1.17) converts the series of scalars into a slightly shorter series of vectors with overlapping entries. In topological terms, each vector can be considered to provide the coordinates of a point in an m-dimensional space. These vectors also provide a way to reproduce the dynamics of the real unknown system.

As Takens (1980) has proven, the trajectory generated from (1.17.), that is, the behaviour of the m-histories constructed from the real data, will mimic the behaviour of the unknown trajectory $\{x_t\}$ generated by the underlying (unknown) dynamical system, provided that m is large enough.

Actually, the evolution of (1.17) creates an image of the real d-dimensional attractor of the system which evolves in the state space R^n, into an m-dimensional space R^m, provided that this latter space, called the ***embedding space,*** is large enough to preserve the dynamical characteristics of the real system.

Following Whitney's "embedding theorem"[10], Takens show that the dimensionality of the embedding space (i.e. the embedding dimension) should satisfy the relationship:

$$m \geq 2d+1 \qquad (1.18)$$

The procedure that creates a "replica" of the unknown attractor of the original system in a lower dimensional space where the attractor can be analysed is called an ***embedding***. Creating an embedding by the use of the delayed copies of a single observable variable is the essence of the method of delays.

The theoretical underpinnings of the embedding procedure are based on principals from differential topology concerning the preservation of the qualitative characteristics between two spaces (vector fields) formally expressed as ***differential equivalence*** or ***diffeomorphism*** [Broomhead & King (1986a)].

Mathematically, diffeomorphism can be defined as a differentiable map (or function) with differentiable inverse or a C^k bijection. Practically, diffeomorphism should be understood as an invertible transport function which takes orbits from one space (or a vector field) into orbits to another space through a smooth change in coordinates. It does so in such a way to preserve their orientation and all their dynamical characteristics. In visual representation terms, the two plots of the orbits in each space will differ only by local stretching while all the qualitative and quantitative characteristics of the flows will remain unchanged.

---

10 This is a basic topological theorem proved by Whitney in 1936 stating that every d-dimensional compact manifold embeds in R^{2d+1}.

Formally an ebedding can be defined as a smooth map, say Φ, from a d-dimensional manifold A (i.e. the original attractor of the system) placed in the state space R^n, to a space R^m, such that its image $\Phi(A) \subset R^m$ is a smooth sub-manifold of R^m and Φ is a diffeomorphism between A and $\Phi(A)$. That is, the embedding of A in R^m is a realisation of A into a sub-manifold $\Phi(A)$ within R^m, provided that the dimensionality of the embedding space is large enough to contain it.

In empirical terms, this abstraction of the real attractor from its phase space and its reconstruction in a lower space by the use of a single dynamical variable makes it possible to extract information about the real (unknown) system by observing the reconstructed one.

Takens (1980) proved that the reconstructed attractor has dynamic properties equivalent to the real (unknown) attractor, that is, identical invariant measures such as entropy, various notions of dimensions and Lyapunov exponents.

Hence, a strange attractor will exist in phase space if a strange attractor exists in state space, both having the same dynamical characteristics, which can be measured by the use of m-histories of real data. In this way, the method of delays, provides a tool to extract hidden information about the unknown dynamics of a system by the use of real observations. Once the embedding space trajectory has been constructed from the m-histories vectors, measures such as dimensions and Lyapunov exponents can be estimated in order to tell if the system has a chaotic deterministic explanation (a strange attractor).

The reconstruction of the phase space of a system is not a straightforward procedure and a bad reconstruction is likely once the parametres involved have not been chosen properly. The two basic parametres to be defined in order to create an embedding are the delay time and the embedding dimension. Their choice is closely related to the calculation of specific measures such as the correlation dimension. Therefore, the embedding procedure should not be considered as an autonomous part of the empirical applications framework, but as a first step of an integrated procedure.

1.6.1.1. *The choice of the delay time (τ).*

Detection of chaos is a time scaling issue. Quantification of chaos is measured by the invariant properties of the attractor (e.g. correlation dimension, Lyapunov exponents etc.). The latter is a reconstructed geometrical object the quality of which depends on the reconstruction parametres. One of the most crucial parametres related to time scaling is the delay time.

The delay time is a positive integer number which determines the sampling rate of the data (the measurable variable) in case that the experiment is not a controlled one, that is, when we cannot interfere to the digitisation process.

According to Gershenfeld (1987), when the data is finite, a bad choice of the delay time might restrict the phase space that is sampled and distort the trajectory constructed by the m-histories in such a way that the original attractor is not properly reconstructed.

Specifically, a very short delay time may create highly correlated m-histories thus causing the trajectory to lie near the diagonal of the embedding space. On the other hand, a large delay time might cause the accumulation of noise and mask the dynamical relationship between the points on the attractor. Ideally, τ should be large enough to maximise the separation between nearby trajectories in phase space, while in the same time it must be as small as possible, in order to avoid contamination by noise.

In empirical terms, the delay time τ must be the shortest time over which there are clearly measurable variations in the observable signal. There are several suggestions in the literature about how this rule can be implemented.

Gershenfeld (1987) suggests a trial and error procedure to select the delay time according to which the attractor of the system is reconstructed and plotted for a variety of different values of the delay time. He claims that in this way a familiarity with the nature of the attractor is gained, which might be proven sufficient for the correct selection of the delay time. This is a rather general consideration, which might be operational when the system to be analysed is a priori known.

Fraser (1989a,b) and Fraser & Swinney (1986) have developed the ***mutual information criterion*** method, which draws heavily on Shannon's [Shannon and Weaver (1963)] information theory. This method uses a box-counting algorithm to calculate the mutual information for a range of small but increasing τ values. The right τ is the one at which the mutual information first reaches its asymptotic value, which is independent from τ. This method, though theoretically appealing, suffers from a high computing cost due to the use of the box-counting algorithm.

An alternative method quoted in Wolf et.al. (1985) specifies the delay time as a function of the ***Mean Orbital Period*** of the dynamical system under investigation according to the relationship:

$$\tau = Q/m \qquad (1.19)$$

where Q is the mean orbital period and m is the embedding dimension.

The mean orbital period or cycle is related to the recurrent character of the trajectories generated by dynamical systems and is not always well defined even for systems with a-priori known equations of motion in the Natural Sciences' empirical framework.

A more efficient method is the ***Autocorrelation Function Decay Criterion***, [Ramsey & Yuan (1989)], according to which τ is chosen to be the time of the first zero in the autocorrelation function (ACF) of the data series. Alternatively, for long autocorrelated series the characteristic decorrelation time (i.e. the minimum time for the ACF to decay to e^{-1}) can be used as the proper delay time [Barnett & Chen (1988)].

This method tries to account for the possible autocorrelation between the data points, which may distort the constructed trajectory. This method is not useful when there are inherent periodicities to the system causing the ACF to occilate for many periods. However, its inherent simplicity makes it the most popular in the literature and is the one employed by this research, as well.

1.6.1.2. The choice of the embedding dimension (m)

The choice of the embedding dimension m is closely related to the calculation of the correlation dimension. Choosing the embedding dimension is in most part of the empirical literature a "trial and error" procedure where successive increasing embedding dimension values are used to create embeddings until the system is properly embedded.

An idea of what "properly" means has already been given by the bound expressed in equation (1.18). However, since the dimension of the real attractor of the system is unknown, this bound is not empirically useful.

The underlying idea to this "trial and error" procedure is that the measurable properties (the invariant measures) of an attractor are not changing after the point where the embedding space is large enough to contain its full shape (its dynamics). So, a d-dimensional attractor will retain its dimension in a higher dimensional space. Consequently, if we create different embeddings for higher embedding dimensions and measure the dimension in each case, once the proper embedding dimension has been reached, the dimension will no further change. This is how the choice of the embedding dimension is related to the dimension calculation since the "proper" embedding can be determined only after this calculation. More explicit details on this procedure are presented in Chapter 2, where the dimension measuring procedure is described.

Another much more empirical alternative to select the embedding dimension is by "guessing" it. This may not look as a scientific approach and it is not in the usual sense. However, the Chaotic framework is primarily (at least for the moment), an experimental field where all kinds of approaches are applied. Yet, it should be clarified that this "guess" is not what it seems. This is due to the fact that there are certain empirical constraints that make it an operational one.

To be more specific, the unknown attractor should be a low-dimensional one in order to be measurable. As we shall see in the sequel, in experimental conditions the dimensionality of a system that can be reliably measured is closely related to data availability. Even for extremely long series (i.e. of many thousands of observations), observable attractors should have low-dimensions, which practically means that embedding dimensions exceeding 8-10 should be avoided [El-Gamal (1987), Smith (1988), Vassilicos (1990)]. Provided that a low dimensional attractor exists, we can have a proper embedding by considering an embedding dimension close to the above limit, even if our system can be embedded in a lower space.

In empirical terms, the latter procedure is used as a first step in the context of the embedding procedure. This is so because the researcher always tries to create the best possible embedding and to this end many different embeddings are always constructed.

1.7. THE SINGULAR VALUE DECOMPOSITION METHOD

An alternative method to reconstruct the phase space of an (assumed) deterministic dynamical system suggested by Broomhead and King (1986 a,b) (hereafter B&K), is the Singular Spectrum Analysis (SSA), also known as Singular Systems Analysis, Singular Value Decomposition (SVD), Principal Component Analysis (PCA) and Karhunen-Loeve Decomposition. Some of these methods are already known from matrix algebra (PCA) or Signal Processing Theory (SSA or Karhunen - Loeve Decomposition) and in general, they provide an algorithm for decomposing multidimensional data into linearly independent coordinates.

The essence of the SVD method is to develop a better way for the projection of the reconstucted attractor of the system, i.e. to generate a new projection basis (coordinate system) for the trajectory matrix **X**, which is an (Nxm) matrix constructed from the delayed vectors of a series as:

$$\mathbf{X} = N^{(-1/2)} \begin{pmatrix} \chi_1^T \\ \vdots \\ \chi_N^T \end{pmatrix} \tag{1.20}$$

where $N^{(-1/2)}$ is a normalisation factor (N being the number of the delayed vectors constructed) m is the embedding dimension, T defines a transposed vector and $\chi_i = \{\chi_1, \chi_{i+1}, \ldots, \chi_{i+(m-1)\tau}\}$, where the components are the same as the componenets of the vector defined in (1.17).

This can be done efficiently by diagonalizing the covariance matrix **V**, which is an (mxm) matrix whose elements are the covariances of the observations forming the χ_i $\{ \chi_i \in R^m \mid i=1,2,\ldots,N\}$vectors:

$$\mathbf{V} = \mathbf{X}^T\mathbf{X} = \frac{1}{N}\sum_{i=1}^{N} \chi_i \chi_i^T \tag{1.21}$$

The covariance matrix can be diagonalized and decomposed to :

$$\mathbf{V} = \mathbf{C}\boldsymbol{\Sigma}^2\mathbf{C}^T \tag{1.22}$$

or equivalently to :

$$\mathbf{V}\mathbf{C} = \mathbf{C}\boldsymbol{\Sigma}^2 \tag{1.23}$$

where $\mathbf{C}$ is an orthogonal m×m matrix whose columns consist of the eigenvectors $\mathbf{c}_i$, $\mathbf{C} = (\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_m)$, and Σ^2 is a diagonal matrix of the eigenvalues $\Sigma^2 = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2)$, assuming that σ_i^2 are ordered from the largest to the smallest $(\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_m^2 \geq 0)$.

Using the definition of $\mathbf{V} = \mathbf{X}^T\mathbf{X}$, equation (1.23) can be further transformed to give:

$$(\mathbf{XC})^T(\mathbf{XC}) = \Sigma^2 \qquad (1.24)$$

The matrix $\mathbf{Y} = \mathbf{XC}$ (an N×m matrix) represents the set of projections of the points of the attractor onto the basis $\mathbf{c}_i$ of the eigenvectors, i.e. it provides the coordinates for plotting the attractor. The components $\mathbf{y}_i$ of the $\mathbf{Y}$ matrix called the principal components (or ***singular functions***), defined also as:

$$\mathbf{y}_i = \mathbf{X}^T \cdot \mathbf{c}_i \qquad (1.25)^{11}$$

represent the coordinate transformation $\{\mathbf{X}^T\mathbf{c}_1, \mathbf{X}^T\mathbf{c}_2, \ldots, \mathbf{X}^T\mathbf{c}_m\}$ that will produce a picture of the attractor. Note that the eigenvectors of $\mathbf{V}$ and the square roots of the eigenvalues of $\mathbf{V}$ are called the ***singular vectors*** and the ***singular values*** of $\mathbf{X}$ respectively.

Accordingly, the set of m eigenvalues $\{\sigma_i^2\}$ is called the ***singular spectrum*** and each σ_i^2 is the mean square projection of the trajectory onto the corresponding $\mathbf{c}_i$.. So, heuristically, the set $\{\mathbf{c}_i, \sigma_i\}$, may be viewed as defining the principal axes of an m-dimensional ellipsoid which is explored by the trajectory and describes, on average, the bounds of the attractor. The $\mathbf{c}_i$ vectors correspond to the directions of the principal axes, while the associated to them eigenvalues σ_i^2 correspond to the lengths of the axes.

SVD seems to have certain advantages over the Takens' method of delays. According to B&K, the latter may introduce linearly dependent coordinates and artificial symmetries into the phase portrait of the system, which increase with the embedding dimension. These symmetries are due to the arbitrary manner that Takens' MOD method uses to choose a basis for the embedding space. On the other hand, the coordinate transformation on delay reconstruction provided by the SVD eliminates these problems. Medio (1992) shows that the choice of $\{\mathbf{c}_i\}$ as a basis for the projection of the trajectory is optimal since the columns of the trajectory matrix become independent and the mean square error of projection is minimised, which means that the artificial symmetry is also eliminated.

In addition, a significant advantage of deriving a projection basis by diagonalising the covariance matrix $\mathbf{V}$ is that orthogonality in the embedding space is related to the statistical properties of the time series. Hence, SVD is actually a statistical method (while Takens' method of delays is not).

1.7.1. The SVD Analysis in Noisy Cases and Noise Filtering

A very useful feature of the SVD method is its ability to filter away noise from experimental data. In the presence of noise[12] the embedding space of a system can be divided into a deterministic and a stochastic (noisy) subspace. Noise makes the singular values of the trajectory matrix to be non-zero, that is, the trajectory seems to explore all dimensions of the embedding space. It has been shown [Medio (1992)] that in the case of a noise contaminated system, the singular value spectrum arranged in descending order will normally show a few "emerging" singular values (corresponding to the principal components) and a flat tail of small positive singular values i.e. a plateau, corresponding to the "noise floor". The singular values lying on the plateau represent noise-dominated coordinates.

Formally, the trajectory matrix $\mathbf{X}$ of a process can be written as:

$$\mathbf{X} = \mathbf{X}_d + \mathbf{X}_s \qquad (1.26),$$

where $\mathbf{X}_d$ and $\mathbf{X}_s$ are the trajectory matrices corresponding to the deterministic and the noise component of the process respectively. $\mathbf{X}_d$ can be expressed in terms of quantities that can be derived from the diagonalisation of the covariance matrix $\mathbf{V}$ as :

$$\mathbf{X}_d = \sum_{i=1}^{d} \mathbf{X}\mathbf{c}_i\mathbf{c}_i^T \qquad (1.27)$$

11 Where (·) is the notation for the internal multiple

12 In most cases noise is assumed to be "white" or "strict white noise", i.e. IID additive noise.

where $\mathbf{X}_d$ is an N×m matrix and d the number of the "emerging" singular values above the noise floor. The "emerging" singular values belong to the deterministic part of the partitioned singular spectrum :

$$\Sigma = \Sigma_d + \Sigma_s \tag{1.28}$$

where: $$\Sigma_d = \mathrm{diag}(\sigma_1, \ldots, \sigma_d, 0, \ldots, 0) \tag{1.29}$$

and $$\Sigma_s = \mathrm{diag}(0, \ldots, 0, \sigma_{d+1}, \ldots \sigma_m) \tag{1.30}$$

The "reduced" (noise free or filtered) trajectory matrix $\mathbf{Y}_d$ can be obtained from the projection of $\mathbf{X}$ onto the reduced singular vector basis that corresponds to the "emerging" singular values as:

$$\mathbf{Y}_d = \mathbf{X}_d \mathbf{C} \tag{1.31}$$

as a (N×d) matrix the rows of which are of the form $\{\mathbf{X}^T\mathbf{c}_1, \mathbf{X}^T\mathbf{c}_2, \ldots, \mathbf{X}^T\mathbf{c}_d\}$, providing a coordinate transformation which produces a picture of the trajectory confined this time to the deterministic subspace of the system.

Gibson et. al. (1992), show that in this case too, the SVD method provides the optimal linear coordinate transformation in the sense that the subsets of principal components have maximum variance and maximum signal to noise ratio. It is also shown that the height of the plateau indicates the variance of the noise (provided that it is "White" noise).

Another important aspect of the SVD is that it reduces dimensionality to an eigenvalue problem. According to B&K, in the case of noise free data, the dimension d of the deterministic subspace coincides with the number of the linearly independent vectors in the embedding space that can be constructed from the trajectory, i.e. to the rank of **V.** Therefore, the number d of positive eigenvalues of the correlation matrix **V** gives the dimension of the attractor while the other (m-d) eigenvalues are equal to zero. In the case of noisy data, Vautard and Ghil (1989) have shown that both the rank of the matrix **V** as well as the level of the noise floor depend on the embedding dimension m and the delay time τ. So they suggest that the number of emerging eigenvalues could be viewed to be giving the "statistical dimension" (SD) of the dynamical system, i.e. a reliable upper bound for the minimum number of the degrees of freedom necessary to describe the data.

In all, the SVD method can be very useful in extracting as much reliable information as possible from noisy time series, for the dynamics of which no prior knowledge exists. According to Gibson et. al. (1992), it provides a crude but robust approximation to strange attractors. It also provides useful information from short noisy series with no prior knowledge about the generating equations. On the other hand, it is essentially a linear method and this has been a point of criticism [Mees, Rapp and Jennings (1987) and Fraser (1989a)]. However, the data-adaptive character of the eigenelements it is based on, makes it valuable in nonlinear dynamics analysis. Another important feature of the method is that it is not data consuming [Vautard et. al. (1992), Palus & Dvorak (1992)]. When truly nonlinear information about low dimensional systems requires many thousands of points, SVD provides useful insight on the dynamics of a system with a few hundred data points. This feature combined with its lack of sensitivity to stationarity [Vautard et. al. (1992), Tong (1992)], makes it a valuable tool in analysing financial series.

Despite its attractiveness and promising characteristics, it should be noticed that SVD is still quite a new method, which has not been widely employed in empirical applications in order to be properly evaluated. So this new approach should not be considered sufficient for understanding all there is to know about a particular time record and works best when it is done in concert with other independent techniques.

Finally, it should be noticed that the SVD method has been used for phase space reconstruction and noise filtering purposes in the Natural Sciences' field [Pilgram et.al. (1992)], but it will be the first time, to our best knowledge, that it will be used with real financial data in the empirical part of this study.

Chapter 2

THE CHAOTIC TESTING[1] FRAMEWORK

2.1. INTRODUCTION

The major empirical question of our concern is whether certain financial series exhibit chaotic behaviour, or alternatively, whether these series are generated by a low-dimensional deterministic system. The first step in the process of investigating an unknown system, assumed to be chaotic, by observing a single variable, is the reconstruction of its phase-space as described in the previous Chapter.
The second step is to employ suitable methods and techniques that can detect chaoticity. In general, the chaotic nature of a system can be detected in two ways:

- An indirect one, which is related to the low-dimensionality of a system, a necessary but not sufficient condition for chaoticity.
- A direct one, which stems from the definition of a chaotic system as a system with Sensitive Dependence on Initial Conditions.

There are two major methods related to each one of the above approaches, namely the Correlation Dimension estimation and the Largest Lyapunov Exponent estimation respectively. Theoretically, the application of these methods should be enough to verify the existence of low-dimensional chaos in a time series representing the unknown dynamical system. However, as we shall see in what follows, the correct application of these methods is a complicated procedure and the use of a wide range of additional techniques is required in order to maximise the reliability of the empirical findings.

---

1 It should be clearly stated that the term "testing framework" is not used here with its proper statistical content since up to now no statistical test for chaos exists. For the same reason we avoid the use of the term "tests" for each of the different methods and techniques employed in this framework.

2.2. THE CORRELATION DIMENSION ESTIMATION

Let a time series $\{x_t\}$, $t=1,..,N_T$, which is assumed to be generated by an orbit that is dense on an d-dimensional attractor, i.e. it represents a chaotic dynamical system that we wish to investigate.

We first create an embedding, i.e. a vector series of m-histories, $\{\mathbf{x}_t^m\} = \{x_t, x_{t+\tau}, \ldots, x_{t+(m-1)\tau}\}$, for a given delay time τ and embedding dimension m. Then the correlation integral (or correlation sum) is calculated as:

$$C_m(r) = \frac{2}{N^2 - N} \sum_{i<j} H_r(\mathbf{x}_i^m, \mathbf{x}_j^m), \qquad 1 \le i \le N,\ 1 \le j \le N, \tag{2.1}$$

where : $H_r(\mathbf{x}_i^m, \mathbf{x}_j^m)$ is the heavyside function that equals one if $|\mathbf{x}_i^m - \mathbf{x}_j^m| < r$ or zero otherwise, $|\,.\,|$ denotes a distance norm, $N = N_T-(m-1)$, and r = the tolerance distance (scaling parameter).

The correlation integral gives a measure of the correlation of points along an orbit, by measuring the fraction of the total number of pairs of m-histories that are within a distance r from each other. If r is increased, each m-history will gain new neighbours (m-histories), the total number of which depends on the character of the underlying process. For a uniform stochastic iid process the correlation integral will expand like r^m, i.e. a scaling law will hold:

$$C_m(r) \sim r^m \tag{2.2}$$

As Brock (1986) and Brock & Dechert (1987,1990) show, in the case of a low dimensional process, the correlation integral will expand independently of m and provided that m is large enough (m>2d+1), it will scale like:

$$C_m(r) = r^{d_m} \tag{2.3}$$

and the correlation dimension can be defined as:

$$d_m = \lim_{r \to 0} \ln C_m(r)/\ln(r) \tag{2.4}$$

Empirically, we embed the system under study in increasingly higher dimensions in order to ensure that the condition $m \ge 2d+1$ is fulfilled. This is because only then d_m is independent of m and is equal to the dimension d of the attractor ($d_m=d$).

For each embedding dimension m, the correlation dimension d_m is calculated, and a plot of the correlation dimension d_m versus the embedding dimension m is created. If the system is a stochastic one, d_m will increase like m. If the system is dominated by deterministic dynamics, d_m will have a finite value (will saturate) after a certain m, which is the "proper" embedding dimension. The saturation level of d_m will give the correlation dimension of the system. A finite but also small value for d_m implies that the system is dominated by low-dimensional chaos, i.e. it is governed by the properties of a strange attractor.

The procedure to calculate d_m for each m, after a vector of m-histories has been constructed, can be accomplished as follows:

First, the correlation integral has to be computed for different values of the scaling factor r and a log-log plot of $C_m(r)$ versus r must be constructed [the G-P plot].

The scaling law will hold over a range of r corresponding to the linear segment[2] of the log-log plot and the correlation exponent is given by the slope of a line fitted (by an OLS method) to this linear segment.

Empirically, we need to specify:

- A norm in order to measure the distance between the pairs of the m-vectors.
- An r region $[r_{max}, r_{min}]$, over which a scaling law should hold.

The distance norm $|\,.\,|$ choice creates no problems. Brock (1986) provides a theorem according to which the correlation dimension is independent of the choice of norm. In the empirical literature both the “Euclidean” and the “max” norms have been used, the latter having the advantage of making computation easier.

The r (or scaling) region choice requires more elaboration. In the case of a set with a finite but large number of points, r should be smaller than the diameter of the set (which is considered to be equal to one) and larger than the mean nearest neighbour distance [Smith (1988)]. Hence, r takes values between 0 and 1 ($0<r<1$). In addition, [see equation (2.4)] r should be small. However, at the limits of very large and very small r, the scaling of the correlation sum is violated. For r large (close to the diameter

---

2 It has been shown that the actual log-log relationship is highly non-linear [Guckenheimer,(1984), Caputo et al.(1986)], but the linear approximation holds for a narrow range of values of the scaling variable r.

of the attractor), the correlation integral saturates at $C(r)=1$ ($\ln C_m(r) = 0$). At the limits of very small r, the violation is due to the finiteness of the data and the presence of noise. Specifically, if r is smaller than the mean nearest neighbour distance, no points or only a few points might be found within r of each other and $C_m(r)$ tends to zero. For r larger than the mean nearest neighbour distance but still very small, any noise present in the data dominates.

In practice there are different ways that the "stable" r region is chosen :

a. From the G-P plot by "eyeballing" [Brock (1986)]. According to this approach, the correlation sum is estimated for each embedding dimension m and for a range of r-values. Usually one or few decades of r values are used and the standard deviation (std) of the data set is in many cases a reference value for r. Then the G-P plot [$\ln C_m(r)$ vs. $\ln(r)$] is constructed and a linear segment of this plot is chosen by "eyeballing". This linear region corresponds to a specific range of r values [r_{min},r_{max}.] for which the scaling low holds. The correlation exponent, for each m, is estimated as the slope of the line fitted by an OLS method to this linear region. Finally, the correlation dimension is given by the saturation level of the plot of d_m versus m, if such a saturation level can be found.

This is a standard procedure in the economic literature and has been followed by most of the empirical works in this field [e.g. Frank, Gencay and Stengos (1988c), Scheinkman & LeBaron (1988,1989), Peters (1991a)]. The major disadvantage of this method is the difficulty in defining accurately the plateau region from the G-P plot.

b. From the plateau region of the plot of the derivative $d'_m = \Delta\ln C_m(r)/\Delta\ln(r)$ versus $\ln(r)$ [Smith (1992)]. In empirical terms the derivative can be defined as :

$$SC_m = [\ln C_m(r_i)-\ln C_m(r_{i-1})]/[\ln(r_i)-\ln(r_{i-1})] \qquad (2.5),$$

which can be considered as a local estimate of the slope of the G-P plot for adjacent values of r.

The "plateau" region of the slope SC_m vs. $\ln(r)$ plot, corresponds to the linear region of the former method giving the r range for which the scaling low holds.

Once the slope/ln(r) plot has been constructed, the correlation dimension can be derived in different ways:

- By regressing (OLS) $\ln C_m(r)$ on $\ln(r)$ for the r range corresponding to the "plateau" region, for each embedding dimension. Thus, the d_m/m plot can be constructed and the correct correlation dimension can be estimated, as it has been already described. This approach has been followed by Barnett & Chen(1988).

- By creating the slope/ln(r) plot for different m values. If the system is a low dimensional one, the curves for each m will converge to the same "plateau" (above some m), at the level of which (on the slope's axis), the correct value of the correlation dimension is found. This approach is often met in the natural sciences' empirical literature [Bountis et al.(1993), Pavlos et. al. (1992a,b,1994)].

  Both approaches should give the same results provided that the "plateau" region is well defined.

This second method is followed by most of the empirical studies in the natural sciences field and is more accurate than the former. It is also the one adopted in this study.

Finally, it should be noticed that some researchers in the economics field use other, yet much more inaccurate methods to estimate dimension. For example, Brock & Sayers (1988) use a local dimension measure estimate, and Frank & Stengos (1988a) estimate the correlation dimension as the arithmetic average of the three highest values of SC_m.

2.2.1. Limitations and Pitfalls.

Dimension algorithms are sensitive to both the qualitative and quantitative characteristics of the underlying system. Stationarity, data length, noise and temporal correlations seem to play an important role in the correct estimation of the correlation dimension.

2.2.1.1 Stationarity

Stationarity is clearly implied by the preceding theoretical analysis in the non-linear dynamics context. Time averages are space averages (as $N \rightarrow \infty$) over the stationary distribution only when the assumption of stationarity is valid, giving us the ability to study the dynamical properties of a system by performing time series analysis [Brock and Sayers (1988), Baek and Brock (1988)].

Non-stationary stochastic series might give a low saturating dimension and thus distort our results with respect to the nature of the underlying process. In practice, when dealing with data from Economics, we face a severe trade-off problem between the requirement for more data and stationarity of the underlying distribution, especially when this data is of medium to low frequency (daily, monthly). In this case, the time spanned is longer and structural shifts are more likely to occur. The stationarity issue is rarely addressed in the Economics' empirical literature, indicating a serious omission in the application of the correlation dimension and the other tests and techniques employed. Notice however, that Barnett and Hinish (1991) claim that solution paths generated by chaotic systems need not be stationary. This, empirically, means that the prewhitening processes followed by many researchers in economics might be doubtful.

In our case, log-differencing of index price series reduces considerably the non-stationarity problem usually met with economic data. In addition, the BDS test is employed to identify whether nonlinearity in our data is due to structural shifts.

2.2.1.2 Data requirements

Small data sets can cause serious bias effects in the correlation dimension estimates. Several empirical studies [Smith (1988), Krishna-Mohan and Subba-Rao (1989), Ramsey and Yuan (1990)] report downward bias in the correlation dimension for random systems and upward bias for chaotic systems when the data is small[3].

As a general rule, the larger the data is, the more accurate the results are. However, the data adequacy question has been a major topic of debate in the literature. Several suggestions exist regarding the data points needed to calculate the correlation dimension.

Gershenfeld (1987) and Mayer-Kress (1987) argue that the adequate number of data points might range from 10^{d} to 100^{d}, where d is the correlation dimension. Other researchers are more specific. Through a complicated formula, Smith (1988) suggests a lower bound of points $N_{min} = 42^{m}$, where m is the embedding dimension. More conservative lower data limits are provided by Nerenberg and Essex (1990) and Tsonis et. al. (1992) [$N_{min} = 10^{2+0.4d}$], Theiler (1990a) [$N_{min} = 5^{d}$] and Eckman and Ruelle (1992) [$N_{min} > 10^{d/2}$].

3 In Ramsey and Yuan (1990), bias effects are also related to embedding dimensions. Bias increases with higher embedding dimension and decreases with larger sample size.

In a more empirical context Grassberger & Procaccia (1983b) suggest that a few thousand points might suffice in the case of low-dimensional (d<3) systems. Abraham et. al. (1986), argue that reliable dimension estimation is possible with small data sets with less than 1000 points in some cases. Ramsey and Yuan (1990) specify a lower limit of 5000-7,000 points for low-dimensional attractors. Vassilicos (1990) finds empirically that 10,000 points are enough for the application of the G-P algorithm to tick-by-tick exchange rate data. Finally Peters (1991a), in a quite different approach claims that it is the time length (in particular the cycles covered by the data) which is important, while the number of points available is of much less importance.

Our empirical results seem to support the more moderate suggestions on the issue [e.g. those of Eckman and Ruelle (1992)] which make possible the application of chaotic techniques to data sets with more than 3000 points.

2.2.1.3 Sensitivity to temporal correlations

Correlation dimension is sensitive to temporal correlations. Theiler (1986,1990a,b) shows that the correlation sum turns out to be biased towards too small dimensions when the pairs entering the sum are not statistically independent. For time series with nonzero autocorrelations independence cannot be assumed. In this case the embedding vectors at successive times may be also close in phase space due to the continuous time evolution. However, determinism should reflect only spatial (geometrical) and not temporal correlations. Theiler (1991) argues that it is more than likely that the majority of the dimension estimates published for field measurements are too low because they mistake temporal coherence for geometrical structure.

The remedy suggested [Theiler (1986)] is a modified version of the correlation integral, given by (2.1) as:

$$C(W,N,r) = \frac{2}{(N+1-W)(N-W)} \sum_{n=W}^{N-1} \sum_{i=0}^{N-1-n} H(r - |\mathbf{x}_i - \mathbf{x}_{i+n}|) \qquad (2.6)$$

where $N = N_T-(m-1)$, N_T = # of points in the time series, m = embedding dimension, H(.) the heavyside function, r the minimum distance allowed between points and $|\ .\ |$ the distance norm.

This modification aims at reducing the temporal correlations between nearby points and bases the correlation dimension estimates on the spatial correlations only. In technical terms, this specification can be considered as a generalised form of the G-P formula. For W=1 the two formulas become equivalent. For W>1, the algorithm does not take into account distances between points that are closer together in time than W, that is, it is functioning as a cut-off parameter to the r value (r>W) [4]. With respect to the choice of W, Theiler notes that, as long as $W>p(2/N)^{2/m}$, (m=embedding dimension), the exact choice of W is not important. Theiler's modification has been applied for the first time to financial data in this study to ensure that our correlation dimension estimates are not biased due to temporal dependence.

2.2.1.4 Sensitivity to Noise

Broadly speaking noise can be defined as any unwanted input either because it is irrelevant or because it distorts the presentation of values and/or their interpretation. Real life systems are noisy in the sense that they are open to exogenous influences that alter their behaviour and contaminate measurements. A major distinction can be done between measurement (or additive) noise and dynamical noise. In the first case the dynamics of a system satisfy $x_{n+1}=F(x_n)$ but we measure scalars $s_n=s(x_n)+\eta_n$, where s(x) is a smooth function that maps points on the attractor to real numbers and $\{\eta_n\}$ is the measurement noise. Dynamical noise in contrast is a feedback process wherein the system is perturbed by a small random amount at each time step: $x_{n+1}=F(x_n+\eta_n)$. Dynamical and measurement noise may not be distinguishable a posteriori based on the data only and in general dynamical noise induces much greater problems in data processing than does additive noise.

The correlation dimension algorithm is sensitive to noise and the case is no better for other metrics such as Lyapunov exponents, entropies etc. In the presence of noise self-similarity (scaling behaviour) is broken and a phase-space reconstruction appears as high-dimensional [Brock and Dechert (1987)] so the correlation dimension estimation may not be able to distinguish between a noisy low-dimensional chaotic process (noisy chaos) and a stochastic alternative. However, it has been empirically shown [Atten et. al.(1984), Theiler (1990a), Smith (1992)] that at a relatively high signal to noise ratios we can still measure the correct correlation dimension of a noisy system.

A serious empirical question is weather noise can be separated from the clean signal. This depends on the nature of the noise. When the noise is additive, noise filtering is

---

4 All x_i,x_j pairs with $|i-j|<W$ are thrown out of the calculation of the correlation integral.

possible but in most empirical situations we cannot identify the source of noise for an unknown system. In Physics literature, various linear and non-linear noise filtering procedures [Kostelish and Yorke (1990), Kostelish (1992), Sauer (1992), Broomhead et al.(1992)] have been suggested, yet, no method has been widely accepted and some of them suffer from serious shortcomings. As Badii et. al.(1988), Mitschke et.al.(1988) and Badii and Politi (1989) have shown, application of certain filters to time series change the geometrical characteristics of the underlying dynamical system and may give false dimension results[5]. A promising alternative for noise filtering is the recently developed SVD analysis, presented earlier, that will be employed by this study.

2.2.1.5 Sensitivity to various types of (autocorrelated) stochastic processes.

Brock (1986) and Brock and Sayers (1988), have shown that near unit root processes, i.e. linear stochastic AR(q) models that can adequately describe a time series, can give low dimension estimates for a large range of r values. This problem is usually met with macroeconomic series since many of them are near unit root processes [Nelson and Ploser (1982)]. In the above case, the G-P algorithm can exhibit the correct scaling for stochastic processes ($C_m = r^m$), but only for very small values of r, which for many cases of small finite data are not workable. Hence, autoregressive processes when dealing with real data can indeed inhibit the ability of the correlation dimension method to distinguish the true underlying process, although a careful application of the test (appropriate choice of the r region) might solve this problem. On the other hand, Sheinkman and LeBaron (1989) show that simulated ARCH processes give high dimension estimates.

Another category of stochastic processes that might fool the G-P algorithm are fractal stochastic processes called "coloured" random noise processes and among them a certain class of random paths such as fractional Brownian motion (biased random walks), which is of interest in the literature of Finance.

Osborne and Provenzale (1989), analyse a process having a power spectrum which exhibits $1/f^a$ scaling over a wide range of frequencies (f) and can be empirically constructed by inverting power-law spectra and randomise Fourier phases. They show that simulated random paths generated this way have low and non-integer finite

5 The filtering issue and the related problems will be further discussed in Chapter 3.

correlation dimension saturating at d=2/(a-1). They argue that this is due to their fractal nature that fools the dimension algorithm. On the contrary, Theiler (1991) claims that it is not the fractality of the time series but the long time correlation that is important, combined with the recurrence time of the trajectory. He argues that the G-P algorithm for small r values and long enough data is able to detect coloured noise processes despite the fact that $1/f^{a}$ noise does not fulfil two major requirements of this algorithm, namely stationarity and lack of autocorrelation. Two more types of stationary correlated noises having power-law spectra ($1/f^{2}$) have been investigated by Provenzale et. al. (1992) who show that they can also fool the correlation algorithm which in this case too gives falsely low estimates.

In the empirical context the most commonly applied method in the Economics' literature to guard against autocorrelation (unit root or AR(q) processes) is the ***"residuals" method*** developed by Brock (1986). He proved that if a time series has a deterministic explanation, then the residuals of a linear time series model with a finite number of lags fitted to the data have the same dimension and the same Largest Lyapunov Exponent (LLE) as the original series. Empirically, the correlation dimension (or the LLE) is first computed for the original series $\{x_t\}$. Then any autocorrelation is removed by the best fitting AR(q) model to the original series and the dimension is calculated again for the residuals $\{\varepsilon_t\}$ of this model. If the original series has a deterministic explanation then the estimates of original series should be the same as those from the residuals. If the original series, although stochastic, give a low and saturating dimension due to autocorrelation, the estimates from the non-correlated residual series will be larger and not saturating, revealing the true (stochastic) nature of the system. Apart from not being a statistical test, the most serious shortcoming of this method is that it may misidentify chaos as randomness when the data is small [Brock (1986)] or/and when the order q of the AR(q) process fitted to the original series is of high order (practically more than 2 or 3) [Brock and Sayers (1988)]. This latter shortcoming limits the power of the method against fractal noise processes that can be very long autocorrelated. A more effective alternative is the "Surrogate data method" presented next.

2.2.2. The Surrogate Data Method & the Bootstrap

If linear autocorrelated or other stochastic processes can give misleading correlation dimension estimates, we need to rule out the possibility that our data could be explained by such a process. This can be done by creating surrogate series and trying to reject the hypothesis that the class of systems represented by the surrogates gave rise to the original series.

Our analysis should be on more solid ground by developing a statistical significance test. Statistical inference or hypothesis testing usually requires a null hypothesis, a test statistic and some means to generate the probability distribution of the test statistic under the assumption that the null hypothesis is true. Parametric statistics derive this probability distribution analytically, however, by invoking restrictive assumptions like Normality. When no parametric statistics are available for non-conventional test statistics, as in our case, a resampling technique like the ***bootstrap*** may be used introduced by Efron in his Annals of Statistics (1979) paper.

It has been shown [Nelson and Startz (1990a,b)] that the bootstrap (under mild regularity conditions) yields an approximation to the distribution of an estimator or test statistic that is at least as accurate as the approximation obtained from 1st order asymptotic theory[6]. So, it can be used to substitute computation for analytical solutions, especially if obtaining the asymptotic distribution of an estimator or statistic is difficult.

The idea behind bootstrapping is simple. Let $X = (X_1, X_2, \ldots X_n)$ a hypothetical random sample from the null hypothesis population with a completely unspecified probability distribution F, and t(X) the test statistic for this sample. Let also $x = (x_1, x_2, \ldots x_n)$ the real sample from the real population and t(x) the test statistic for that real sample. A hypothesis test can be formulated as :

$$\text{Reject the null if } prob\{t(X) \geq t(x)\} \leq \alpha. \qquad (2.7)$$

Thus, the problem in assessing a significance level reduces to estimating the sampling distribution of t(X). If the null hypothesis population can be completely specified, this

---

6 Actually, the bootstrap is often times more accurate in finite samples than 1st order approximations. Horowitz (1997) provides good insight of the bootstrap's ability to improve upon 1st order asymptotic approximations. This is the case when the bootstrap is applied to statistics that are "asymptotically pivotal", i.e. to statistics whose asymptotic distribution is independent of unknown population parameters. Bootstrap estimates of the distributions of statistics that are not asymptotically pivotal have the same accuracy as 1st order asymptotic approximations.

estimation can be done by Monte Carlo sampling which, however, requires the standard assumptions made in conventional parametric tests. Alternatively, according to Efron (1979), the sampling distribution of t(X) can be estimated on the basis of the observed data x as follows :

First the empirical distribution of the observed data, i.e. the sample probability distribution $\hat{F}$, is constructed placing probability mass 1/n at each observation $x_1,x_2,\ldots x_n$. This probability function is the maximum likelihood nonparametric estimator of the population distribution [Efron and Tibshirani (1993)].

In a second step, with $\hat{F}$ fixed, a surrogate data set, i.e. a random sample of size n, is drawn with replacement from $\hat{F}$, which is called the bootstrap sample $_{B}x$.

Finally, the sampling distribution of the test statistic $t(_{B}x)$ is calculated by generating repeated realizations of $_{B}x$. The bootstrap conjecture is that the bootstrapped sampling distribution of $t(_{B}x)$ approximates the sampling distribution of t(X) if the sample contains all the available information about the population ($F=\hat{F}$).

There are parametric and non-parametric implementations of the bootstrap. For instance, in a regression context one typically bootstraps OLS residuals. The non-parametric implementation of such a bootstrap scheme would be to bootstrap directly from the residuals (or a weighted version of them). The parametric implementation of this scheme would be to sample from a normal distribution with mean zero and variance equal to the estimated variance [see Cribari-Neto and Zarkos(1998)]. There are also alternative ways to apply bootstrapping. For example, instead of using random sampling of single observations with replacement for the bootstrap, blockwise (resampling with replacement moving blocks of consecutive observations) bootstrap construction can be used [Kunsch (1989), Liu and Singh (1992)]. In all cases stationarity of the original sample is an implied assumption for bootstrapping [Jeong and Maddala (1993)].

The major advantages of the method are that the null hypothesis population does not have to be defined since the sample serves as its proxy and that a significance level can be estimated for any test statistic even when its exact sampling distribution has not been derived.

Its principal disadvantage is that the validity of the bootstrap conjecture cannot always be assessed and depends upon the form of the test statistic. Assessing the power of the bootstrap method is feasible only when parametric statistics are also available for the

test statistic. For those cases there is good evidence that simple test statistics such as standard error is expected to bootstrap successfully on the basis of about 50 - 250 samples, but longer series of bootstrap samples (500-5000) might be needed when hypothesis testing is the objective [Lunneborg (1987)]. The problem is that the null hypothesis may be rejected for the wrong reason, i.e. because the shape of the sampling distribution is not well approximated by the shape of the bootstrap sampling distribution. In addition, as Efron (1979) notices, even when bootstrap works well, fundamental inference problems remain due to the fact that the bootstrap provides approximate frequency statements and not approximate likelihood statements.
However, the bootstrap method remains a valuable tool when conventional parametric tests do not exist as in most of the methods employed in this study.

In our case the fundamental issue is to increase our knowledge about the nature of the process (system) that gives rise to the observed data. This can be done by creating surrogate series and test the hypothesis that this class of processes gave rise to our data. Of course the increase of our knowledge depends on the hypothesis made and whether or not it is rejected.
This type of hypothesis testing should be able to discriminate between the original series and the artificially created surrogates. Each surrogate represents a different null hypothesis process *p,* postulated to be the source of the original data and for stochastic processes *p* is a probabilistic law. By comparing the statistical properties of the surrogate data with those of the actual series will help us to draw conclusions about the system under study. What we actually test is how a test statistic for the actual data is compared to the sampling distribution of this statistic for *p*.
We saw earlier that one method of generating surrogate data in the bootstrap is by sampling with replacement from the empirical distribution. Alternatively we can obtain surrogate samples corresponding to different null hypotheses from transformation of the original data by the use of several techniques discussed in the next paragraph. In our application of the bootstrap methodology we have followed the steps described below:

1. A discriminating statistic is selected according to the method employed and its value g_0 is estimated for the original data. Notice that for most of the test statistics in this study (e.g. correlation dimension, Hurst exponent) no parametric statistical

framework exists.

2. Depending on the hypothesis we want to test against, we create a bootstrap sample $_{B}p$ from the original series by employing alternative techniques. The bootstrap is comparable to the original data in certain respects but is also consistent with the null hypothesis we are testing for. For example, by randomising our original series we can create a surrogate that is random iid but has the same distributional characteristics as the original series.
3. A number N_B of repeated realisations of $_{B}p$ are created and for each one of them the value of the discriminating statistic is calculated as g_i, $i= 1,\ldots,N_B$. This process gives the bootstrap sampling distribution of the test statistic $g(_{B}p)$. In our example a ensemble of $N_B=1000$ repetitions of the randomisation procedure of our original series are performed and the value of the test statistic g_i, $i=1,\ldots,1000$ is estimated for each one of the surrogate samples.
4. The bootstrap sampling distribution and g_0 are used to assess a significance level and reject or not the null hypothesis. Obviously a low significance level indicating that g_0 is a rare value for $g(_{B}p)$ means that the null hypothesis should be rejected. In our example again (when the Hurst exponent is used as the test statistic), if $\text{prob}\{g(_{B}p)\geq g_0\}\leq 5\%$, then the null hypothesis that our data might consist of independent random numbers drawn from a fixed but unknown distribution can be rejected at the 95% significance level.

Further details and the exact form of the null hypothesis to be tested in each case will be discussed separately for each application in the following Chapters.

Regarding the possibilities of a successful bootstrap in our applications we should mention first the implied stationarity assumption. Recall that the stationarity issue as the implied hypothesis for the applications employed in the nonlinear dynamics framework and in this study specifically has been already discussed.

A second issue is the adequate number of the bootstrap samples' repeated realisations in order to have a good approximation of the sampling distribution of the test statistic. The existing literature provides no specific guidance on this issue and as already mentioned the range of the suggested realisations is very wide ($50<N_B<5000$). As it will be shown in each separate application we have used both the upper and the lower limits of this range depending on how difficult it was to calculate the test statistic.

Finally, a more general issue is the difficulty in assessing the power of the method due to the lack of a parametric statistical framework for the test statistics we employ. However, this is an unavoidable compromise until the appearance of a statistical test for chaos.

2.2.2.1 Alternative techniques for constructing the surrogate series

Empirically, the surrogate data method entails the creation of surrogate series, which are comparable to the observed data in certain respects to be specified, but which are also consistent with the null hypothesis we are testing for. This can be done in different ways.

In the Economics' literature, two different approaches of this method have been applied in the context of the correlation dimension estimation. The first is referred to as the ***"wing"*** diagnostic while the second, and most commonly applied, as the ***"shuffle"*** diagnostic.

In the first case a Gaussian surrogate is constructed having the same length, mean and variance with the original series. The name "wing" diagnostic [Brock and Sayers (1988)] is due to the shape of the GP plot [$\ln C_m(r)$ versus $\ln(r)$] that can be alternatively used as a graphical representation of the test. If the series are non-Gaussian then the "wing" shaped GP plot for the Gaussian random series will lie below the respective GP plot for the original data and vice versa.

In the second case Gaussianity is replaced by the less restrictive hypothesis of iid-ness. The underlying idea is [Scheinkman and LeBaron (1988)] that if a chaotic structure exists in the original series, a shuffling procedure will destroy it and the dimensionality of the new series will rise. Empirically, the original series are treated like an "urn" and by sampling from them randomly with or without replacement we create iid surrogates. In the latter approach, which is the one followed by this study, this randomisation process produces surrogate iid series, having the same length and distributional characteristics as the original data.

It should be noticed that in almost all cases of empirical applications of the above mentioned techniques in the Economics literature no statistical hypothesis testing has been performed. Instead, a single surrogate series has been constructed and the correlation dimension for the surrogate was simply compared with the respective estimate for the original series.

Another approach to create surrogates is the ***"Phase Randomisation"*** technique developed in the Natural Sciences' empirical framework. The first simplified version of this approach is due to Grassberger (1986) but more recent applications can be found in Theiler (1986,1991) and Provenzale et.al. (1992). The usefulness of this technique stems from its effectiveness against linear correlation, or more general, against fractal noise processes (e.g. like the $1/f^{a}$ noise). In addition, according to Provenzale et.al.(1992), it may be helpful against non-linear (multifractal) stochastic processes which can also exhibit low-dimensionality.

In this case the surrogate data is created by randomising the phases of the Fourier transform of the original data. Specifically, in order to test against the null of linearly correlated noise, random surrogate series having the same length, variance and autocorrelation as the original series, are created. This is done by taking the Fourier transform of the original series given by:

$$x(t_i) = \sum_k C_k \cos(\omega_k t_i + \varphi_k) \qquad (2.8)$$

where (φ_k) are the Fourier phases of the series $x(t_i)$ and C_k are constants, randomising the phases and then taking the inverse Fourier transform. Different choices of the random phases will create several sets of surrogate data in order to quantify the statistical significance of our results by bootstrapping, as mentioned above. In the case of stochastic processes, the saturation of the scaling exponent in the correlation dimension estimation is forced by the shape of the power spectrum and this is consistent to the fact that both the power spectrum and the correlation integral are related to the second moments of the distribution. On the contrary, for a chaotic signal phase correlations play an essential role. Thus, the invariance of the correlation dimension (or any other statistic) under phase randomisation, i.e. the invariance of the correlation dimension between the original and the randomised series, implies that the convergence of the correlation dimension in the original series is determined only by the shape of the spectrum (equivalently by the autocorrelation function). This in turn means that the null of linearly correlated noise cannot be rejected.

In the case that the original series is a chaotic one, the dimensionality of the surrogate series will show (via bootstrap) significant differences between the original and the phase randomised series. Nevertheless, a change in the correlation integrals must be interpreted with caution when we deal with periodic or quasi-periodic processes against which the power of the phase randomisation test is limited.

2.3. THE LARGEST LYAPUNOV EXPONENT ESTIMATION

Several different approaches have been suggested in the literature for the calculation of the Lyapunov Exponents (LE). Some of them apply when the equations of motion of the dynamical system are known and in this case the calculation of the complete Lyapunov spectrum is feasible and relatively easy [Benettin et al. (1980), Shimada and Nagashima (1979)]. In the case of experimental data, computing the LE spectrum is not a straightforward task. However, in this case the computation of the dominant (largest) LE suffices to characterise a chaotic system if this exponent is found to be positive.

Different procedures have also been suggested by the literature for the calculation of the Largest Lyapunov Exponent (LLE). Some of them are limited to one-dimensional maps [Gollub et al. (1980), Hudson & Mankin (1981), Nagashima (1982), Wolf & Swift (1984), Wright (1984)] and are not suitable in the case of higher-dimensional systems.

There are different methods that can deal with higher-dimensional systems [Eckman et. al. (1986)] but the most popular among them is the one suggested by Wolf et al. (1985) that is employed by this study, too.

This method measures the divergence of nearby points in the reconstructed phase space and displays the scaling of the rate of divergence over fixed intervals of time in order to calculate the largest LE. This procedure is anything but straightforward and involves a number of parameters to be specified, a task which depends mainly on experimentalization. Analytically the method can be described in the following steps:

1. After an m-dimensional phase portrait from the original series x(t) has been reconstructed by the method of delays, the first step is to choose two close enough points, considered to lie on neighbouring trajectories. These two points can be considered to represent the initial conditions or the early state of the first principal axis $p_i(0)$ of the ellipsoid [see (1.14) in Chapter 1]. The first point is picked to serve as the central one to be followed (as time evolves) and is considered to lie on a, so called, "fiducial trajectory", that is, the trajectory along which the expansion/stretching (but not the contraction/folding) of the attractor should be measured by the LE. The data record is then scanned to depict the second point, which should be the closest (in the

Euclidean sense) to the first one, but not closer than the noise level (or SCALMIN), which must also be defined. Let this initial distance length be L(t).

The distance between the two points is followed as the central point is stepped through the data set for a fixed interval called the evolution time (EVOLV=n-time steps). After this time the distance between the two points is measured again and let the new distance be L'(t+1). The log of the rate between the initial and the final length [$\log_2(L(t)/L'(t+1))$] gives the rate of divergence between the two initially selected points. If this rate is enormously large, most probably, the initial points do not lie on the same region of the attractor and their initial closeness is due to the way the attractor is folded. In this case a new second point must be selected and the procedure is repeated. If this is not the case, the procedure advances to the next step which is the non-fiducial point replacement phase.

2. The replacement of the non-fiducial point is a necessary step in order to measure only the stretching or divergence in phase space. When the separation between the initial points becomes large, they might fold into one another and the measurement will involve convergence, which is not part of the largest LE. Consequently, the choice of the evolution time is important in order to avoid folding. The choice of the new replacement point should satisfy two criteria: it should preserve the orientation of the original evolved elements (the reference trajectory) and its separation from the evolved fiducial point should be small. That is, both the orientation change (ANGLEMAX) between the original pair of points and the new pair and the replacement length or evolution length (SCALMAX) should be minimised. Empirically, this means that after a fixed evolution time, the data points are scanned in order to find a replacement point the angular error of which should be less than ANGLEMAX=θ, and its distance from the fiducial point should be less than SCALMAX but far from SCALMIN. In the case of more than one candidates, the point with the minimum angular change is chosen. In case that no candidate is found we can either alter the two criteria's values (SCALMAX and ANGLEMAX) or retain the points that were been used.

This replacement procedure is repeated at $t = it_s$, $i = 1,2,\ldots,M$, where t_s is the evolution time (EVOLV) and M is the total number of replacement steps, until the fiducial trajectory transverses the whole data set, by which time we hope to observe a stationary behaviour of the largest LE (L_1). L_1 is calculated as a running average rate of trajectory divergence which is updated after each replacement step :

$$L_1 = \lim_{M \to \infty} \left\{ \frac{1}{Mt_s} \sum_{i=1}^{M} \log_2 \left[\frac{L(t_i)}{L(t_{i-1})} \right] \right\} \tag{2.9}$$

The LLE estimation suffers from the same data requirements problem, in both quality and quantity terms, as the correlation dimension test. According to Wolf et al.(1985), a minimum of 10 orbits of data (orbital periods) is required for stable Lyapunov estimates. The problem is how we can estimate this orbital period with real data. Peters (1991a,1994), provides a solution by suggesting that the mean orbital period in real data can be estimated by the use of Rescaled Range (R/S) analysis, a method that will be employed by this study, as well.

Measurement (additive) noise creates no problem in defining Lyapunov exponents. However, dynamical noise may create problems in the calculation of the Lyapunov exponents, but in this case Wolf et. al. (1985) argue that the system may be characterised by the LE of the noise free system averaged over the range of noise-induced states.

Another difficulty arises from the implementation procedure for the calculation of the Largest Lyapunov Exponent which is not straightforward. As it is obvious from the description above, a lot of parameters have to be defined. Unfortunately it is not clear in the literature how these parameters can be specified, especially in the case of real data. However, some implementation suggestions can be helpful.

The choice of the embedding dimension (m) is important. In general, if m is too low, interleaves between distinct parts of the attractor are likely, leading to biased estimation. If m is too large, noise contamination of the data will occur, tending to decrease the density of the attractor points, thus making the replacement procedure difficult. Moreover, in some cases, increasing m increases the surface curvature of the reconstructed attractor making it difficult to retain the orientation constraints. In practice, the choice of the embedding dimension (m) should be relevant to Takens's rule ($m \geq 2d+1$). However, smaller m values (just greater than the dimension of the attractor) might yield reliable LE estimates, [Takens (1983b), Wolf et al.(1985)].

With respect to the time delay (τ) choice, we can follow the rules described in the case of the correlation dimension test (r value equal to the decorrelation time).

Alternatively, Peters (1991a) and Wolf et. al. (1985), suggest to follow the $Q = m\tau$ rule, where Q is the mean orbital period of the system.

The propagation time (or evolution time-EVOLV) between replacements should be a trade off between: a) too frequent replacements, which are computationally costly and can cause a loss of phase space orientation, and b) too infrequent replacements, which reduce orientation errors and computing cost but may allow folding of the trajectory, when we want to measure only stretching. For systems with unknown dynamics the use of a range of evolution times is recommended in order to check the stability of the LLE [Wolf et. al. (1985)].

The evolution length (SCALMAX) is again a trade off between lower orientation errors with shorter lengths and noise which is higher in shorter vectors. On the other hand, the minimum scale (SCALMIN) should be adjusted to the level below which noise is expected to dominate. In practice, we can either use proportions of the attractor's length to fix these parameters according to the suggestions of Peters (1991a) and Wolf et. al. (1985), which in a series can be translated to proportions of the range of data. SCALMAX is suggested not to exceed 5-10% of this range while SCALMIN is usually fixed to 10% of SCALMAX. Alternatively, Bountis et. al.(1993), suggest that these two parameters should correspond to the points on the X axis defining the "plateau" region in the slope vs. ln(r) plot, used to define the correlation dimension of the system under consideration.
Finally, the angular orientation (ANGLMAX), according to Wolf, is not considered as a free parameter since it was not found to have much effect on exponent estimates and is normally fixed to 0.2-0.6 radians.

2.4. ADDITIONAL METHODS TO IDENTIFY CHAOS

2.4.1. The "Signal Differentiation" Method.

This is another method suggested in the Natural Sciences empirical context [Bountis et. al., (1993), Provenzale et. al., (1992)]. The surrogate data in this case is the first (numerical) derivative or difference, of the original data. If a system is governed by a low-dimensional attractor then the correlation dimension estimate of its first (or higher) derivative should remain unchanged since the dynamics of the original data and the differenced one would be expected to have the same geometrical and statistical properties. Indeed, as Provenzale et.al.(1992) have shown, this is the case for the first differenced signal of the x component of Lorenz attractor.

Conversely, in the case that the data has a significant stochastic component, the correlation dimension of the differenced data, if well defined at all, is expected to be much larger than the respective estimate of the original data. This is, for example, the behaviour exhibited by the two fractal noise signals analysed by Provenzale et. al. (1992), as previously described .

It should be noticed that this test might be proven very useful in the case that data from Economics is employed, especially when the data is suspected for non-stationarity, since the first difference transformation may also account for non-stationarity in the mean[7]. For example, the application of this test could lead to totally different conclusions about the nature of some series reported to be chaotic, as in Peters (1991a), and in Barnett & Chen (1988), who use detrended data instead of log-differenced series.

2.4.2. The "Independent Realisations" method

According to this test [Bountis et. al., (1993), Provenzale et. al., (1992)], several independent realisations of the dynamics of the system under study are considered.

---

7 Notice that the application of the test is useful in the case of stationary data, too. For example, Boundis et. al.(1993) applied it to a clearly stationary surface temperature data which seemed to pass all other tests for chaoticity and found that a significant stochastic component was present .

Empirically, this means that the data is divided into several consecutive pieces, each piece representing an independent realisation which has different initial conditions from the original signal.

In the case that the original signal is a chaotic one, these independent realisations, when taken together to form the original signal, should display the same "geometrical" and statistical features as each one of them taken by itself. That is, the correlation dimension of the original series (one realisation) should be the same to the respective estimate from the superimposition of the independent realisations.

On the other hand, in the case that the system is stochastic, the different realisations tend to fill the entire space and the correlation dimension is increased with the number of the realisations considered. Provenzale et. al.(1992), show again how this test works when applied to data from the Lorenz attractor and to the respective data from their two fractal noise signals.

This method shares the same principal with tests for nonstationarity, employed by this study too, where different subsamples of the same series are tested in order to depict structural changes [Hsieh (1991)]. However, in the context of chaotic analysis very long series are required since the subsamples created should be long enough to be analysed by the chaotic methods. Hence in economic series, the applicability of this test is limited.

It should be noticed that in our empirical analysis to follow, all the above mentioned methods and techniques have been employed for both markets under investigation (Athens and London stock markets) with two exceptions. Firstly, the "signal differentiation" method has not been used since our data consists of log-differenced series. Secondly, the "independent realizations" test has been used with the London Stock Market series only, due to the adequate length of these series.

Chapter 3

NON-LINEAR DYNAMICS & CHAOS IN THE MARKETS

3.1 A THEORETICAL PERSPECTIVE

3.1.1 The traditional (stochastic) framework

The traditional theory explaining the price formation in Financial Markets is the Efficient Market Hypothesis (EMH), which in an empirical context is expressed via the expected returns or the «fair game» efficient market model [Fama (1970a, 1976), Samuelson (1965), Mandelbrot (1963)].

A restrictive form of market efficiency and a special case of the «fair game» model is the well-known Random Walk model assuming independent and identically normally distributed returns (iid). In the early '70s, Random Walk was considered to be the most respectful model in describing price formation in Financial Markets [Godfrey et. al. (1964)]. In the early literature, most of the studies looking for non-zero autocorrelations, seemed to support the weak form hypothesis and Random Walk specifications [Alexander (1961), Cootner (1964), Fama (1965, 1970b), Fisher (1966), Jensen and Bennington (1970), Granger (1972), Malkiel (1981)].

However, several studies have cast doubt on random walk by revealing different kinds of anomalies mainly in the form of seasonal patterns. Among them "Intra-day effect" is reported by Harris (1986), "Day-of –the–week effect" by Cross (1973), Rozeff and Kinney (1976) and Gibbons and Hess (1981), "Weekend effect" by French (1980), Keim and Stambauch (1984) and Connolly (1989), "End-of-month effect" by Gutelkin and Gutelkin (1983) and Ariel (1987), "January effect" by Keim (1985), Haugen and Lakonishok (1988) and Lakonishok and Smidt (1988), "Holiday effect" by Zarowin (1988) and Ariel (1990) and other kind of anomalies [Keim (1983), Kato and Schalheim (1985)].

In more recent research, examination of the validity of the "weak form hypothesis" is mainly concerned to discover any form of dependence in the return series, indicating predictability. Fama (1991) presents the relevant literature as «tests for return predictability». Short-term autocorrelations of returns are reported by Lo and Mackinlay (1988) and Conrad and Kaul (1988) indicating predictability and rejection of the market efficient-constant expected returns model. Long-range dependencies have been found by Fama and French (1988) and Poterba and Summers (1988). Short-term predictability of returns from various variables is also reported. Among the variables used to predict returns are expected inflation [(Bodie (1976), Fama (1981)], short-term interest rates [Fama and Schwert (1977)], dividend yields [Rozeff (1984)] and earnings/price ratios [Campbell and Shiller (1988)].

Additional sources casting doubt on the linear approach and random walk came from the behaviour of volatility, which was found higher than the expected in a rational market's framework [Shiller (1981, 1989)] and unstable through time [Turner and Weigel, (1990)] as well as from the empirical evidence of deviations from normality and iid-ness in return series [Copeland and Weston (1988)].

Based on the above mentioned inconsistencies with a linear view, a number of non-linear stochastic models have been developed [e.g. the Bilinear model by Granger and Anderson (1978), the Threshold Autoregresssive model by Tong and Lim (1980) e.t.c.]. Special notice among them should be given to the Conditional Heteroskedasticity or (G)ARCH-type models [Engle (1982), Bollerslev, (1986,1987a,b)] triggering a voluminous literature during the last few years, an excellent review of which is given by Bollerslev, Chou and Kroner (1992).

Despite their popularity, there is evidence against the assumption that time varying volatility can account for all the existing dependencies. Several studies report evidence of remaining nonlinearity after (G)ARCH filtering [Blank (1991), Vaidyanathan and Krebiel (1992), Yang and Brorsen (1993), Abhyankar et. al. (1994)]. In addition, according to bispectral tests [Hinich and Patterson (1985), Ashley et. al. (1986), Ashley and Patterson (1989)] GARCH processes generate zero bispectra, which is inconsistent with the non-zero bispectra evidence in stock returns [Brock and Malliaris, (1989)]. Finally, (G)ARCH-type specifications cannot cope with kinds of structural instability such as the dependence on the investment holding period or the portfolio aggregation level, put forth by some empirical studies [Brock (1987), Hinish and Patterson (1988), Hsieh (1991), Barnett and Chen (1988), Pilarinu (1993)].

3.1.2 The Chaotic alternative

Recently, another non-linear alternative has been brought to the attention of financial researchers, namely chaos theory. Chaos theory is not a new financial theory but a new mathematical tool in the same way that stochastic processes is the mathematical tool used by the current financial theory, i.e. EMH, Capital Market theory, Portfolio theory etc. Chaotic dynamics is a special type of nonlinearity and chaotic models offer a rich variety of behaviour which can describe disorder, discontinuities and erratic behaviour, properties which characterise financial markets.
Chaotic dynamics imply inherently unstable systems. Chaotic specifications can account for departures from normality, independence and rationality, yet unlike their stochastic counterparts they result in disequilibrium market clearing points. In fact chaotic dynamics imply a more general notion of equilibrium consistent with feedback effects and non-stabilising nonlinearities, confined within the attractor set. In chaotic models all sources of the system's dynamics are endogenized. Fluctuations are produced internally and uncertainty is due to the structural instability and sensitive dependence on initial conditions resulting in a stochastic appearance with deterministic origin i.e. to an oxymoron scheme where randomness coexists with order. Thus, chaotic modelling is fundamentally different from the well-known log linear models with Gaussian error terms.

An important question is whether a chaotic alternative could be theoretically feasible in financial markets. Current financial theories are characterised by simplifying notions like stable equilibrium, rational behaviour and homogenous beliefs of the market participants. The basic assumptions common to all the existing financial theories (CAPM, APT, OPT e.t.c.), which lead to stability and equilibrium, are rational behaviour of individuals and common knowledge beliefs among all investors. Traditional financial theory assumes the existence of risk averse rational investors, individually and in the aggregate, requiring mean/variance efficiency. This behavioural framework has been brought into question. Larsen, Mosekilde and Sterman (1990) claim that chaotic dynamics can be produced by human decision making behaviour. At the individual's level different studies [Kahneman and Tversky (1979,1984), Tversky (1990), Tversky and Kahneman (1974,1991), Savit (1989)] argue that irrational and non-linear behaviour of the investor cannot be ruled out, which means that in the aggregate non-linear feedback effects and behaviour based on fads and fashions is possible. Along the same line, Peters (1991a) argues that fat tails in the return distributions might not be due to the arrival of information in clumps, but instead to the nonlinear reaction of the market participants to information[1] in the sense that they do not react to information until a trend has been established. Such a behavioural framework is consistent with chaotic dynamics.

Common knowledge is a strong assumption, which allows for aggregation in the framework of all the major financial models. In case of homogenous beliefs but not common knowledge, i.e in markets with decentralised knowledge, multiple equilibria consistent with complex behaviour, is possible. Pilarinu (1993) and Ploeg (1985) provide different settings under which the lack of common knowledge leads to chaotic regimes. Arrow (1987) points out that instability might be generated by the multiple interactions of ex-ante rationally acting individuals.

Moreover, homogeneity of investors is also questionable and diversity is a more plausible suggestion. Peters (1991a,1994) argues that the same single holding period investment horizon assumed by Capital Market Theory is not the case and investors react differently to information since they have different investment horizons. Brock and Hommes (1997,1998) show also how the hypothesis of heterogeneous beliefs can produce chaotic dynamics from standard financial models.

Hence, if some of the basic assumptions of the current financial theory are relaxed according to the real world market conditions, unstable equilibrium and the possibility of chaotic dynamics cannot be ruled out.

The implications of a chaotic alternative are important and justify the interest in examining the relevance of chaos theory to finance. Firstly, prediction and control in financial markets should be mentioned. Chaotic systems make possible short-term prediction as well as supervisory stabilisation policy [Brock et al. (1991)] while none of the above is possible in the case of stochastic specifications. Moreover, the different sources of uncertainty (unpredictability of future shocks in stochastic models vs.

---

1 Linear reaction to information could lead to linear feedback effects, i.e. to mean reverting behaviour. However, empirical evidence from R/S analysis [Peters (1991a, 1994)] shows trend-reinforcing behaviour.

structural instability and sensitivity to initial conditions in chaotic ones), imply quantitative and qualitative differences in risk measurement.

Regarding market efficiency, we should make clear that chaotic modelling of asset prices and returns is not necessarily incompatible with EMH. A chaotic approach may be partly compatible with EMH, in the sense that prices might fully reflect information through sensitivity to initial conditions. In addition, although chaotic processes are short-term forecastable, this could be possible at horizons too short to allow for profitable exploitation by speculators. On the other hand, in a purely chaotic scenario price movements do not need external shocks and existing equilibrium asset pricing models should be replaced with dynamic ones allowing for different regimes in a bounded «randomness».

In the existing chaotic literature we see that accepted financial theory is possible to be used to construct chaotic models in financial markets. Most of the existing theoretical models that have been developed in the economics literature are trying to incorporate chaotic paths to conventional models with an intertemporal linkage. This literature includes chaotic versions of economic growth models [Stutzer (1980), Pohjola (1985), Day (1982,1983), Dechert (1984), Boldrin (1988)], the overlapping generations model for rational decision making [Benhabib and Day (1980,1981,1982), Grandmont (1985)], business cycle models [Lorenz (1987b)], international trade models [Lorenz (1987a)], as well as, a variety of other specifications [Dana and Motruccio (1986), Deneckere and Pelican (1986), Kelsey (1988), Baumol and Benhabib (1989), Woodford (1989), Boldrin and Woodford (1990), DeCoster and Mitchell (1991)].

In the finance literature, most models that have been developed are rather simplistic. Smith and Hoy (1990) present a chaotic model that can explain a stock market crash, Shaffer (1991) uses a simple model based on the logistic map to show that a firm policy of proportionate dividend payouts can produce chaotic price behaviour.

Larrain (1991a,b) has developed an interest rates model called the K-Z model, which combines a behavioural map based on Keynesian economics (the Z-map) with a chaotic component (the K-map). The model allows for a shift in the investors behaviour (rational or irrational) depending on the prevailing conditions. Vaga (1991) develops a dynamical nonlinear statistical model called the Coherent Market Hypothesis (CMH). The model allows for transitions in price behaviour from random walk to orderly trends and chaotic fluctuations. More recently Brock and Hommes (1998) show how the simple asset pricing model of Lucas (1978) can exhibit an extremely rich asset price dynamics, including chaos under the hypothesis of heterogeneous expectations.

A more fundamental issue is to try to determine definitively whether chaotic behaviour can indeed be observed in financial time series. This different empirical direction should precede the development of chaotic models explaining market behaviour. This direction, which is the one followed by this dissertation, is of a diagnostic nature and employs a testing framework that is capable of distinguishing chaos from randomness. If a new theory has to be developed and new models permitting chaotic dynamics have to replace the existing ones, this must be justified, firstly, by empirical evidence in the markets. This motivation triggered the empirical work in Economics and Finance, the major contributions of which are presented in the next section.

3.2 EMPIRICAL APPLICATIONS IN ECONOMICS

A substantial part of the literature covering the last decade (1988-1998) is concerned with the question of whether time series from Economics and Finance can be, in part at least, characterised by low-dimensional chaos. A quite complete presentation of this literature in the form of summary of published results appears in Table 3.1.

As can be seen from this Table a wide range of different time series has been analysed of low (annual) to very high (tick by tick) frequency and length that ranges from 60 to 100,000 observations. The method most commonly used is the Correlation Dimension estimation (CD) followed by the Largest Lyapunov Exponent estimation (LLE) and the BDS test, both of which are employed by more than half of the studies. Additional tools such as the "residual" method and the "shuffle" diagnostic have also been used by some of the studies in conjunction with the correlation dimension estimation in order to further evaluate the empirical results. Finally a few studies use various other approaches, such as Kolmogorov Entropy, R/S analysis and nonlinear forecasts.

As regards the main conclusions of the literature, there is a broad consensus of support for the proposition that the series analysed are characterised by a pattern of nonlinear dependence. This conclusion is reached in almost all cases by the use of the BDS test, after proper filtering since the BDS cannot directly test for nonlinear dependence [its (rejected) null is iid-ness]. On the contrary, the evidence on low dimensional chaos is mixed. Part of the papers report weak or strong evidence of chaotic structure, part of them are inconclusive and a significant part finds no indications of chaoticity whatsoever.

The mixed results reported might simply be due to the nature of the different series tested. In general, each different series in terms of the variable to be analysed, its origin (country, market etc.) or the time period covered could give different results with respect to the underlying dynamics. However, serious problems affecting the reliability of the empirical findings could also be responsible for the ambiguities in the literature. Taking into account that a statistical test with a chaotic null is still lacking, these problems can arise from the limited range of the testing framework adopted or/and from the improper application of this framework.

As we can see in Table 1, in most cases in the literature the conclusions are based on the application of very few or even of a single test. Given the lack of a statistical

Table 3.1 Empirical Literature – Summary of Published Results

AUTHORS	DATA	SAMPLE INFO.	TESTS FOR NONLINEARITY & CHAOS	EMPIRICAL FINDINGS	CONCLUSIONS REG. CHAOS
Abhyankar et. al. (1994)	High-frequency(15sec.-1min) stock returns (US, UK, J, GER)	2268-97185 obs., (Sept-Nov. 1991)	1) BDS to AR,GARCH residuals 2) LWG test (Lee,White & Granger (1993) to AR residuals 3) LLE	1) Nonlinearity 2) Nonlinearity 3) Negative	No chaos
Abhyankar et. al. (1995)	High-frequency stock returns (UK-FTSE100)	60000 obs.	1) Bispectrum 2) BDS 3) LLE	1) Nonlinearity 2) Nonlinearity 3) Negative	No chaos
Abhyankar et. al. (1997)	High-frequency(15sec.-1min) stock returns (US, UK, J, GER) & Futures (US, UK)	11000-100000 obs.(sept.-Nov. 1991)	1) BDS to AR,GARCH residuals 2) LWG to AR residuals 3) LE-Kolmogorov Entropy	1) Nonlinearity 2) Nonlinearity 3) Negative	No chaos
Alogoskoufis & Stengos (1991)	Annual Unemployment series (US,UK)	100-200 obs. (1892-1987)	1) BDS to ARCH residuals 2 LLE	1) Nonlinearity (UK series) 2) Negative	No chaos
Barnett & Chen (1989)	Weekly Monetary aggregates (US)	800-850 obs log-linear detrended (1969-1984)	1) CD 2) LLE	1) Low cd (d=1.3-1.5) 2) Positive	Chaos
Barnett & Hinish (1991)	Monthly Monetary aggregates (US)	200 obs. Log-differenced (1969-84)	Bispectrum	Nonlinearity	Compatible with chaos
Barkoulas & Travlos (1998)	Daily stock returns (ASE-30)	2200 obs. (1981-90)	1) BDS to AR & ARCH residuals 2) CD to AR & ARCH residuals, shuffle diagn. 3) Kolmogorov entropy	1) Nonlinearity 2) Low cd (d=5.8) but weak saturation, passed the residuals & the shuffle diagnostics 3) Positive & saturating	Incoclusive
Bask (1996)	Daily exhange rate returns	2409 obs. (1986-1995)	1) CD 2) LLE	1) Low cd (d=2.8-6.5) 2) Positive	Chaos in one case, inconclusive for the rest
Bask (1998)	Daily exhange rate returns	380-700 obs. (1991-95)	LLE (Bootstrapping)	Positive and significant	Chaos
Blank (1991)	Daily US Futures returns (commodities)	250-5823 obs. (1966-1987)	1) BDS to GARCH residuals 2) CD, Shuffle diagn. to GARCH residuals 3) LLE to GARCH residuals	1) Nonlinearity 2) Low cd (d=1.3-6.2), passed the shuffle diagn. 3) Positive	Chaos
Brooks (1998)	Daily exhange rate returns (pound based)	5191 obs (1974-94)	1) CD to raw data & Randomized phase surrogates 2) LLE	1) Low cd in most cases (d=0.9-6) 2) Negative	No chaos
Brock & Sayers (1988)	Quarterly US Macroeconomic series	130-150 obs. Detrended (1948-1985)	1) BDS to AR residuals 2) CD to raw data & AR resid. 3) LLE	1) Nonlinearity 2) Low cd (d=2.5-3.5), failed the residual diagn. 3) Positive	Inconclusive

Table 3. 1 (Continued)

Cecen & Erkal (1996)	High frequency exhange rate returns	3191 obs. (Jan-July 1986)	1) BDS 2) CD 3) LLE	1) Nonlinearity 2) Not saturating cd 3) Positive	No chaos
Chyi Yih-Luan (1997)	Daily stock returns from Taiwan	5000 obs (1976-1993)	1) BDS to raw data,GARCH-M & Fuzzy model residuals 2) CD to raw data GARCH-M res.	1) Nonlinearity 2) Low cd (d=2.4-3) for raw data but higher for filtered data	Weak evidence of chaos
De Coster et. al. (1992)	Daily US Futures returns (commodities)	4087-5308 obs. (1968 – 1989)	CD and shuffle diagn. to raw data AR and ARCH residuals.	d = 6-11, passed the shuffle and the residual diagnostics.	Inconclusive
Doran (1990)	Annual Gold price data	60 obs. (1851-1910)	CD	Not saturating	No chaos
Eckman et.al. (1988)	Daily US stock returns	5200 obs.	LLE	No clear evidence of SDIC	Weak evidence of chaos
Frank & Stengos (1988)	Quarterly Canadian macroeconomic series.	160-250 obs. Log-differenced (1947-1986)	CD to raw data and AR residuals, shuffle diagnostic.	Low cd (d=2.4-3.5) for some series that failed the residuals diagn. and passed the shuffle diag.	Inconclusive
Frank & Stengos (1989)	Daily Gold & Silver returns	3000 obs. (1974-1986)	1) CD to raw data and ARCH residuals, shuffle diagnostic. 2) Kolmogorov entropy	1) Low cd (d=6-7) for raw and res. series, passed the shuffle diag 2) Positive and saturating	Chaos
Frank et. al. (1988)	Quarterly GNP (Italy,Japan,UK,Germany)	88-110 obs. Log-differenced (1960-1987)	1) BDS 2) CD to raw data and AR resid. Shuffle diagnostic 3) LLE	1) Nonlinearity (Japan GNP) 2) Low cd (d=2.1), passed the residuals and shuffle diagnostics (Japan GNP only) 3) Negative	No chaos
Hsieh (1991)	Daily & Weekly US stock returns	1250-2017 obs. (1963-1987)	1) BDS to AR residuals 2) 3rd moment test 3) Locally Weighted Regression	1) Nonlinearity 2) Not nonlinear in-the-mean 3) No forecast inprovement	No chaos
Linden et. al. (1992)	Daily/ weekly UK Stock ret.	1100-2700 obs. (1970-90)	CD	Varying cd (d=7-11)	Inconclusive
Liu et. al. (1992)	Daily US Stock returns	5903 obs. (1962-1985)	CD	Not saturating cd	No chaos
Mayfield & Mizrach (1992)	High frequency (20 sec.) US Stock returns	19027 obs. (Jan. 1987)	1) CD and shuffle to raw data, AR & GARCH residuals. 2) Kolmogorov entropy	1) Low cd (d=3.5) and all tests passed when the method of delays is not used. 2) Positive & saturat.	No chaos when the method of delays is used.
Medio (1992)	Daily exchange rate returns	4204 obs. (1973-1989)	1) CD 2) LLE	1) Low cd (d=2.1) 2) Positive	Chaos
Panas & Ninni (1998)	Daily oil products price returns from 2 markets	1161 obs. For each series (1988-1993)	1) CD to raw data & GARCH res. 2) Kolmogorov Entropy 3) LLE	1) Low cd (d=3.6-6.2) 2) Positive and saturating 3) Positive	Chaos for 10 out of 14 series

Table 3.1 (continued)

Peters (1990)	Monthly stock returns/ prices (US,UK J, GER.)	370-450 obs.(1950-1988) Prices are log-linear detrended	1) R/S 2) CD, shuffle diagnostic. 3) LLE	1) Persistence, long-term memory 2) Low cd (d=2.3-3.05), shuffle diagn. passed 3) Positive	Chaos
Peters (1994)	Daily US stock returns	24900 obs. (1888-1990)	R/S	Persistence, long-term memory	Compatible with noisy chaos
Scheinkman & LeBaron (1989)	Daily & weekly US stock returns	400-2500 obs	1) BDS to AR residuals 2) CD and shuffle to raw data, AR and ARCH residuals.	1) Nonlinearity 2) Low cd (d=6), residuals & shuffle diagnostics passed	Chaos
Sewel et.al. (1996)	Weekly stock returns (6 markets)	730 obs. (1981-1994)	1) BDS to ARMA and GARCH residuals 2) K- and Z-map models	1) Nonlinearities in some cases 2) Significant K- and Z-map results	Inconclusive
Takala & Viren (1996)	Monthly Finnish macroeconomic series, exchange rates, stock returns	893 obs in each series (1922-1994). All series are prefiltered by AR(4).	1) BDS to ARCH residuals of the prefiltered series 2) CD to the prefiltered series 3) R/S analysis	1) Nonlinearity 2) Not saturating in all but one cases 3) Long-memory effects	No chaos
Timmermann & Satchell (1992)	Daily stock returns (12 markets)	2758-3116 obs. (1980-1992)	1) CD 2) Nonlinear forecasts	Low cd (d=2-3)	Indications of Chaos not economically exploitable
Vaidyanathan & Krebiel (1993)	Daily US futures misprising series	1600 obs. (1987,1988)	1) BDS to AR & ARCH residuals 2) CD	1) Nonlinearity 2) Low cd (d=6.5)	Chaos
Vassilicos (1990)	High frequency exchange rates ($/DM)	4584-20408 obs. (9-15/4/1989)	CD	Not saturating cd	No chaos
Tata & Vassilicos (1991)	High frequency exchange rates, daily US stock returns	5000-29137 obs.(1885-1988)	1) CD, shuffle 2) LLE	1) Not saturating cd 2) Zero	No chaos
Vassilicos et. al. (1992)	High frequency exchange rates ($/DM)	20000 obs. (4/89-6/89)	Capacity dimension	Multifractality	Inconclusive
Yang & Brorsen (1992)	Commodity returns (9 series)	2348-2524 obs.(1979-1988)	1) BDS 2) CD to raw data and GARCH residuals	1) Nonlinearity 2) Low cd (d==0.3-0.21) for two series but no saturating cd after GARCH filtering	No chaos
Yang & Brorsen (1993)	Daily US Futures returns (commodities)	2500 obs. (1979-1988)	1) BDS to raw & GARCH resid. 2) CD to raw and GARCH resid.	1) Nonlinearity 2) No clear saturation of cd	Inconclusive

chaotic framework, this can easily give totally misleading results. In general, the number of the methods and techniques used in the literature is limited, especially when compared to the arsenal of the Natural Sciences. From the "surrogate series" based techniques only the "shuffle" diagnostic is used in a few cases. The Bootstrap method which in conjunction with the surrogate method can provide a statistical framework for applications such as the CD and the R/S test, are not used, either. Important contributions such as Theiler's (1986) modification for effectively removing temporal dependence are totally missing.

However, a more serious shortcoming of the literature is the improper or, in some cases, the naïve application of the methods employed, which may lead to unreliable or false results. The calculation of the CD and the LLE with 60-200 observations only [in Doran (1990), Alogoskoufis and Stengos (1991), Brock and Sayers (1988) and Frank et. al (1988)] are typical examples of naïve application of these methods. On the same "data-length" basis Vassilicos (1990) criticises the results reported by Barnett and Chen (1988) and Scheinkman and LeBaron (1989), arguing that the saturation of the correlation dimension might be a small data effect.
The results of these two studies are also criticised by Ramsey et. al. (1990) who replicate them. They argue that the low-dimension estimates reported by Scheinkman and LeBaron might be due to the nonstationarity in their data and not to an underlying chaotic structure. In the case of Barnett and Chen's work, their criticism concerns the sample construction (by spline interpolation), which creates smoothing effects, biasing the dimension estimates downward.

In general, the empirical literature is characterised by non-uniformity at the application level of each test. Many technical details differ (the algorithms used and the embedding and other application parameters vary among the different studies) and is difficult to assess the impact of this non-uniformity on the reported results. We should stress however the absence, in most of the studies, of the delay-time parameter in phase-space reconstruction, a standard procedure in the Natural Sciences testing framework. How serious this omission can be and how reported results might be jeopardised is shown in Mayfied and Mizrach (1992).

On the other hand, in most of the studies from the Economics field the methodology followed has some similar characteristics. The data is firstly pre-treated by some detrending or log-differencing procedure as an adjustment for non-stationarity. Then the series are filtered for any stochastic dependence (usually by AR or (G)ARCH specifications) and the testing framework is applied to the residual series.

In the Natural Sciences literature the approach is different. When experimental noise-free data is used, no pre-treatment is necessary. When the series is noisy, noise-filtering techniques are used to clear the series and to isolate the structured part, which is identified by the use of different methods. The above differences are related to a serious methodological problem, which Barnett and Hinish (1991) call the "Maintained hypothesis" problem (MH).

The empirical question underlying to the MH problem is whether chaos should be the null hypothesis to be tested or whether the maintained hypothesis should be extended to include linear or non-linear stochastic processes, as well. Due to the lack of a statistical test for chaos, this problem is not resolvable at the moment. However, Barnett and Hinish argue that the approach followed by the Natural Sciences is the appropriate one to be applied to Economics literature too, since it is more relevant to the classical statistical methodology. That is, chaos should be considered as the deterministic null hypothesis, which is permitted to explain as much of the variability in the data as possible under the null, while the residuals are assumed to be random (high dimensional noise). On the other hand, exhaustive pre-processing leaves no room for chaos since there will always be a stochastic process [e.g. AR(1), AR(2) or AR(459)] to explain part of the data's variability.

It is true that the methodology followed by most of the studies in the Economics field suffers from serious shortcomings. Pre-treatment of data through detrending procedures may impose trend reversions [Nelson and Ploser (1982)] that give misleadingly low dimension estimates as Frank and Stengos (1988a) and Frank et. al. (1988) show. In Peters (1991a) log-linear detrending of his stock price data generates a series with high autoccorelation against which no precautions are taken. So the low dimension results reported are highly questionable. In all, detrending might give false indications of a chaotic structure when no such structure is present.

On the other hand, filtering for stochastic dependence might create the opposite problem. Theiler and Eubank (1993) stress the ill effects of bleaching the data and show how it affects upwards the CD estimates. Fitting stochastic linear or nonlinear models to remove any stochastic dependence might wipe out part of the existing chaotic structure and diminish the power of the tests and methods for detecting nonlinearity and chaos. Brock and Sayers (1988) and Sheinkman and LeBaron (1989) raise this issue by showing how the residuals of high order AR and ARCH models give misleadingly high CD estimates rejecting the hypothesis of chaos too often when it is true.

The above discussion reveals the complexity of the problem of adapting chaotic analysis in real financial series. In the same time, it justifies the scope of this research aiming, in the empirical level, at the adoption of a methodology capable in minimising the above mentioned problems and shortcomings, thus ensuring the empirical findings, as much as possible. The principal is to employ a wide range of the existing tools and to be extremely careful in incorporating and fully exploiting the most recent theoretical and empirical findings, wherever they come from. We use most of the approaches adopted in the Economics field and a serious effort has been made to avoid the above mentioned problems and shortcomings. However, where we focus on is the adoption of the methodology and the extensive use of the methods and techniques from the Natural Sciences field, most of which are applied for the first time to financial data.

The above mentioned approach is what we call a **"multiple testing methodology"** and in the following Chapters we shall show how efficient it can be in analysing a time series in a non-linear dynamics framework in the absence of a traditional statistical framework.

3.3 DATA AND METHODOLOGY

The two markets under investigation are the Athens Stock Exchange (ASE) and London Stock Exchange (LSE), represented by their official Indices, which are not adjusted for dividends.

Our Greek data consists of the first differences of the natural logarithms of daily Index prices of Athens Stock Exchange (ASE) which give returns in continuous time. The Index is the ASE official one, being broad based and value-weighted. The data covers a 13-year period from 1/1/1981 to 31/10/1993 consisting of 3181 observations being the largest ever used in Greece. In addition it is an original data per-se, since it has been collected by a piece-by piece time consuming procedure, due to the lack of computational facilities in Athens Stock Exchange. This manual collection procedure raises a data validation issue irrespectively of how careful this piece-by-piece data collection procedure has been. To this end, an effort was made to identify alternative sources that might have gathered the same data or at least part of it for their own use. Fortunately two such sources were identified, a Mutual Funds company (KTIMATIKI AEDAK) and a stock broker firm (OMEGA S.A.). The data from both the above sources was received in electronic format and was compared to our data. The first of these sets covers the period from 1/83 – 5/94 while the second from 6/80 to 12/92. The differences found from this comparison were no more than 1% of the total data set. For each particular observation found to differ between the data sets a second search was done in the original archives of ASE and the correct value was established. This way our whole data set has been validated. It should also be noticed that before the submission of this thesis, ASE started to use information technology extensively and historical data was also available in electronic format. So, we had the opportunity to cross check our data set once more and this time no mistake or omission was found.

Our British data consists of the first differences of the natural logarithms of daily prices of the value-weighted Financial Times All Share Index for the period from January 1969 to June 1994 and has been collected by the use of the DATASTREAM facilities at the University. This data set consists of 6653 daily returns, one of the largest samples in daily frequency that has ever been used in empirical studies.

Validation of the British data was proven to be very important and helped decisively to avoid serious flaws in our analysis and results. Specifically, our original sample spanned the period from 1/1/65 to 30/6/94 consisting of 7693 observations. Notice that data prior to this date was not available through DATASTREAM. To avoid possible mistakes during the downloading procedure, the whole data set was downloaded twice on different dates. As an additional validation measure, monthly and weekly data was downloaded too and was compared to the corresponding observations of our daily data set. However, during the visual inspection of the observations (a procedure that has been followed for the Greek data, too), we noticed that from 1/65 to 12/68, daily observations within each week were artificially interpolated. Table 1 in the Appendix shows the price data for the first two months of 1965. It is obvious from this table that only weekly (Wednesday) observations are the real ones and the rest are interpolated values with a fixed step of increase (0.2). This naïve interpolation induces artificial structure to the series that may produce totally misleading results. This is shown in Figure 1 (a-g) in the Appendix where the time series plot the phase-space plot and the correlation dimension estimate are presented for the LSE series from 1/65 – 12/68 (1042 observations) and for a second sample from 1/69 – 12/72 (1043 observations). The comparison of the three plots for the two samples (that in order to be completely comparable have the same length), reveals a clear artificial symmetry for the first, which also leads to a very low correlation dimension estimate. As it will be shown in the following Chapters the behaviour of the total series (1/69 – 6/94) that has been used as our final sample is also completely different from the behaviour of the first 1042 "artificial" observations that have been left out. Notice that should the total series had been used, test results would be biased[2] and conclusions would be seriously flawed.

Methodologically, a statistical description of the data is first performed in the time and frequency domains. The data is then tested for departures from independence (iid-ness) and, after linear filtering, for nonlinearity, by the use of the BDS test since non-

---

2 Although we don't report the results here, correlation dimension estimates are found to saturate at a rather low dimension ($d \cong 5$) when the total data set is used. This is a biased result due to the artificially constructed part of the series that could lead to the conclusion that chaos is present in the LSE data.

iid and nonlinear data are consistent with a chaotic structure. The BDS test is also employed to test our data for non-stationarity.

In the next step long term-memory and fractality of the data is investigated by the use of Rescale Range (R/S) analysis, which is also a useful tool in distinguishing between fractal noise and chaos. We show that classical R/S analysis can be a powerful tool when combined with the bootstrap method for assessing the statistical significance of our results. As an additional test for long-term memory and an efficient tool to account for autocorrelation, Lo's modified R/S statistic is also applied to our data.

The nonlinear dynamics analysis which follows, includes:

(i) Visual inspection techniques which include the time series plot, the phase-space plots and return-maps. These techniques are rarely useful in the case of noisy data series, however, we use them to compare our series with purely chaotic and random data and observe any existing differences.

(ii) The Correlation dimension estimation as well as related methods and techniques that serve as additional tools to validate the correlation dimension results. These tools are the "residuals" method, the "wing" and the "shuffle" diagnostics the "phase randomisation" and the "sign randomisation" techniques. The bootstrap method has been employed to provide statistical inference for most of these methods. It should also be noted that for more accurate results in the estimation of the correlation dimension we employ techniques adapted from the Natural Sciences fields (delay time reconstruction, Theiler's-W modification).

(iii) The Lyapunov exponent estimation which, however, is of limited use in distinguishing chaos from randomness, as we shall explicitly show.

Singular Value Decomposition (SVD) is used in the next step, for phase space reconstruction and noise filtering purposes. Noise filtering can be useful in isolating chaotic components that could be masked by noise. Noise-filtered series are tested again for chaotic structure by the use of the R/S analysis, correlation dimension and Lyapunov exponent estimation.

Forecasting techniques are used as additional tests to distinguish chaos from randomness. Two nearest neighbour non-linear forecast methods are used, the piecewise approximation method and the Simplex method. Three different approaches based on these algorithms are employed, the "DVS", the "Varying Prediction Time" the "Dimensionality" techniques.

Finally, different linear and nonlinear methods are used for short-term forecasting. The results are statistically assessed, but more interestingly the economic value of the forecasts is assessed, too, in "almost" real trading conditions.

We claim that, although none of the methods and techniques employed in this study can compensate for the lack of a statistical test for chaos, this framework is one of the most complete in the existing literature, so far. It permits a thorough analysis of the series analysed and due to the complementary character of the multiple methods used, minimises as much as possible the uncertainty of the results.

Chapter 4

TESTS FOR INDEPENDENCE AND FRACTALITY

4.1 STATISTICAL DESCRIPTION

The basic statistical properties of the ASE and the LSE returns are listed in Table 4.1 while in Figure 4.1 a,b the time series plots of the two series are presented (Index prices and returns). In both sets, the excess skewness and leptokurtosis observed are well known characteristics of stock returns [Mandelbrot (1963,1966), Fama (1965)] and indicate in this case, too, a strong departure from the normal distribution. For the ASE series, all moments are more pronounced than their counterparts in the LSE series.

Table 4.1 Statistical description of the daily returns of the ASE and LSE markets.

Statistical measure	ASE returns	LSE returns
Sample size	3181	6650
Average	0.000831	0.000497
Median	0.00016	0.00034
Variance	0.000292	0.000105
Standard deviation	0.017099	0.010239
Minimum	-0.16291	-0.12117
Maximum	0.24227	0.08978
Range	0.40518	0.21095
Lower Quartile	-0.00481	-0.00487
Upper Quartile	0.00559	0.00612
Interquartile range	0.0104	0.01099
Skewness	0.68473	-0.26952
Standardised skewness	15.7662	-8.9729
Kurtosis	25.1199	9.9705
Standardised kurtosis	289.198	165.967

These differences as well as deviation from normality for both series are shown in Figure 4.2 a,b where the distribution of the two series is presented with a fitted normal curve. It is obvious that both series are leptokurtic, left skewed with fat tails, a picture quite common in stock market returns.

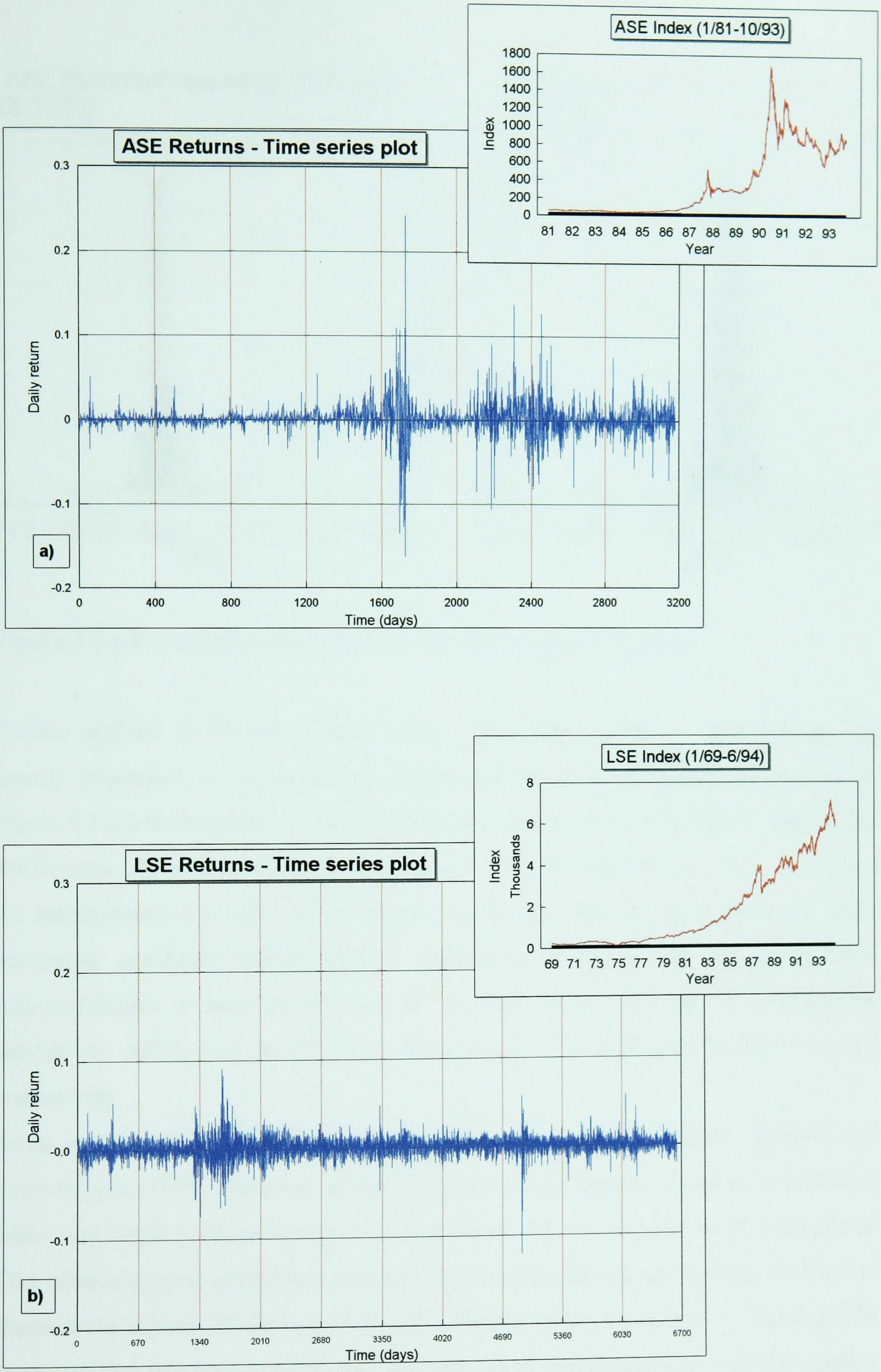

Figure 4.1 a,b Index price (small graph) and Index returns plots of the ASE and the LSE series.

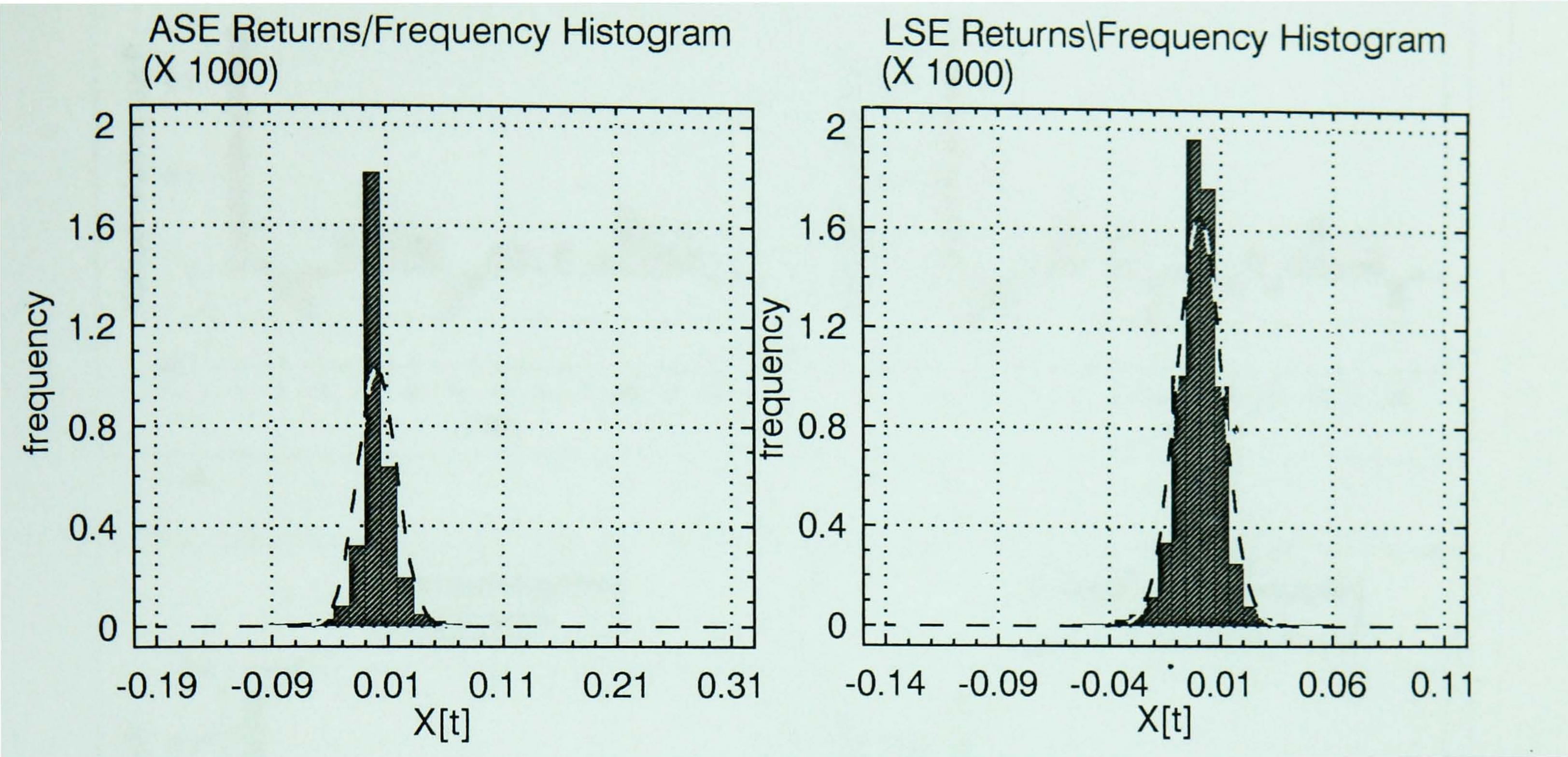

Figure 4.2 a,b Frequency histograms of the ASE and the LSE series

Further analysis in the time domain, shows that daily returns in both markets are linearly dependent, as the autocorrelograms and partial autocorrelograms indicate in Figure 4.3 a,b for the ASE series and in Figure 4.3 c,d for the LSE series. Notice that the boundary lines in the above plots correspond to 1% significance level. As we can see autocorrelation is very strong at least for the first lag, but there are some more marginally significant autocorrelation coefficients in longer lags for both series. Autocorrelation is more pronounced in the ASE series, and partial correlograms indicate an AR(2) and an AR(1) specification for the ASE and the LSE returns, respectively.

Serial correlation is a widespread phenomenon in financial markets. Satchell and Timmermann (1992) find most of the European stock market indices autocorrelated, while very liquid markets like the U.S. and Japan, show also some serial dependence. This phenomenon is probably associated with non-trading (asynchronous trading) of shares in the indices. In the case of the LSE market, which is a liquid one, this could be due to the fact that we use the FTALLSHARE index, including all stocks in the market, some of which might not be traded on a daily basis. In the case of the ASE

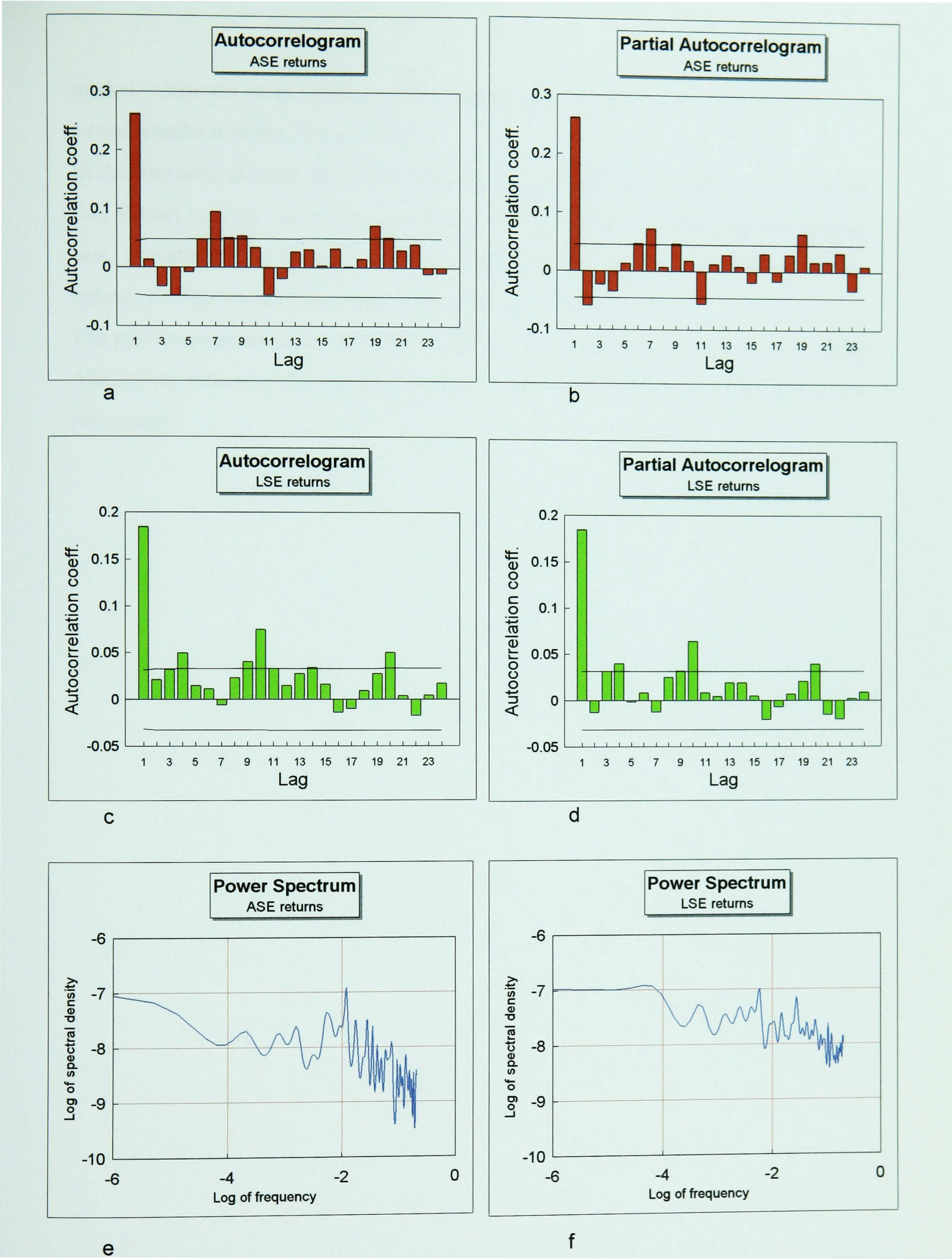

Figure 4.3 a-d : Autocorrelograms and partial autocorrelograms for the ASE and LSE returns.
e-f : Power spectrum of the ASE and the LSE series

market, where the index consists of the most tradable stocks, this is probably due to the overall smaller liquidity of the market[1].

In the frequency domain, the power spectrum of the ASE series (Figure 4.3 e), shows two distinct regions. The left-hand region, which is almost horizontal to the frequency axis, indicates strong presence of noise, while the right-hand one exhibits a decline, indicative of power law behaviour and a non-white noise structure.

The power spectrum of the LSE data (Figure 4.3 f) looks quite similar to that of the ASE series, indicating also strong presence of noise, yet, a power law behaviour is less pronounced.

---

1 It should also be noticed that significant autocorrelations in longer (than the first two) lags in the autocorrelograms might indicate a possible seasonality in variance in the form of day-of-the week effects e.g. weekend effect. However, this kind of conclusion requires more specific analysis since, especially in the case of the ASE series, a lot of non-trading days for various reasons disturb the concequetiveness of the trading days.

4.2 THE BDS TEST FOR INDEPENDENCE

4.2.1 Description of the test

The BDS test, originally suggested by Brock, Dechert & Scheinkman (1987)(hereafter BDS) and Dechert (1988), is a powerful test for independence and is based on the correlation integral concept [Grassberger and Procaccia (1983a)].
According to this test, given a time series $\{x_t: t=1,...,N_T\}$, we can form m-dimensional vectors (m-histories) $\mathbf{x}_t^m = \{x_t, x_{t+1},..,x_{t+m-1}\}$, where m is the embedding dimension, and calculate the correlation integral defined as:

$$C_m(r,N_T) = [2/(N^2-N)] \sum_{i<j} H_r(\mathbf{x}_i^m,\mathbf{x}_j^m), \qquad 1 \leq i \leq N,\ 1 \leq j \leq N \tag{4.1}$$

where : $H_r(\mathbf{x}_i^m,\mathbf{x}_j^m)$ is the heavyside function that equals one if $|\mathbf{x}_i^m-\mathbf{x}_j^m|<r$ or zero otherwise, $|.|$ is the sup-norm, $N = N_T -(m-1)$, and r = tolerance distance.

BDS showed in a statistical context of U and V statistics [Denker and Keller (1986)], that if $\{x_t\}$ is a random series of independent and identically distributed (iid) observations, then :

$$C_m(r,N_T) = C_1(r,N_T)^m, \tag{4.2}$$

and the statistic :

$$B(m,r,N_T) = N_T^{1/2} [C_m(r, N_T) - C_1(r, N_T)^m], \tag{4.3}$$

converges to a normal distribution with zero mean and variance $V(m,r,N_T)$, which can be consistently estimated from the sample data as:

$$V(m,r,N_T) = 4\left[K^m + 2\sum_{k=1}^{m-1} K^{m-k} C^{2k} + (m-1)^2 C^{2m} - m^2 K C^{2m-2}\right] \tag{4.4}$$

$$\text{where, } C = C_1(r, N_T) \text{ and } K = \frac{6}{N(N-1)(N-2)} \sum_{i<s<j} H_r(x_i,x_s) H_r(x_s,x_j) \tag{4.5}$$

Thus the statistic:

$$W(m,r,N_T) = B(m,r,N_T)/[V(m,r,N_T)]^{1/2}, \tag{4.6}$$

known as the BDS statistic, converges to the standard normal distribution N (0,1) and statistical inference is possible.

Intuitively, the test is a formal representation of the difference between the GP plot of the data, reflecting the behaviour of $C_m(r, N_T)$, and the GP plot of a scrambled counterpart which is iid and scales as $C_1(r, N_T)$.

It has been shown by Monte-Carlo simulations, [Hsieh and LeBaron (1988), Hsieh (1991), Barnett et. al. (1992)], that even for quite small samples (N_T =500 obs.), the asymptotic distribution of the BDS statistic approximates iid data from different distributions (e.g. standard normal, student-t, chi-square, Cauchy etc.) well and can detect linear, stochastic non-linear and deterministic non-linear departures from iid-ness, for r values ranging between 0.5 and 2 standard deviations of the data and m values between 2 and 6.

The BDS test, although a test for statistical independence, has been extensively used in the literature to detect nonlinearity, provided that any linear dependence has been removed first from the data. It has been reported [Brock et. al. (1988), Hsieh (1989)] to be superior to other non-linearity tests, such as the Tsay test [Tsay (1986)] and the bispectrum test [Hinish (1982), Hinish and Patterson (1985)]. An important advantage of the BDS test against the former ones is its robustness to moment condition failure in both the asymptotic and the sampling distributions. This is especially important when financial series are tested, since it has been found [Loteran and Phillips (1994), DeLima (1994)] that heavy tailed distributions like stock returns, exhibit moment condition failure, i.e. finite second but non existing fourth moments. Both the Tsay and the bispectum tests are particularly sensitive to the problem of moment condition failure [DeLima (1994)].

In the empirical analysis to follow we use the BDS test to detect nonlinearity in our data, since nonlinearity is consistent with a chaotic explanation. We also use it as a tool to investigate the possibility of nonstationarity.

4.2.2 Empirical evidence

Our series have been filtered first, by a best-fit time series linear model in order to remove any linear dependence. To select the appropriate linear model we have used the information derived from the partial autocorrelograms (Figures 4.3 b and d) as well as Schwarz's information criterion [Schwartz (1978)].
Both these criteria indicate the proper AR lag for the linear model to be fitted and in the case that a different lag is indicated by each method we chose the longer one.
In our case both criteria indicate an AR(2) specification for the ASE series and an AR(1) specification for the LSE series.

Table 4.2 shows the BDS statistic for the original (unfiltered) ASE series, the AR(2) residuals (filtered series), the randomly shuffled AR(2) (or scrambled) residuals as well as for Gaussian random data, having the same mean and variance with our original series. This random data will also be used research for benchmarking purposes in other occasions in this research.
The randomly shuffled residuals serve as an additional control measure to the reliability of the BDS, since shuffling will destroy the supposed non-linear structure in the data and the BDS statistic will capture the difference between the shuffled and the unshuffled series.

The BDS test results in Table 4.2 strongly suggest that daily ASE returns are nonlinearly dependent. The test statistic is significant at the 1% level (the critical value being 2.576 for the standard normal distribution) in both cases of the original and filtered data under all embedding dimensions and r sizes that have been used, thus rejecting the iid-null hypothesis. Notice that filtering reduces the value of the W statistic value, indicating that existing linear dependence affects indeed the BDS results. Accordingly, the BDS statistic is clearly non-significant (at the same 1% significance level) in both cases of the scrambled residuals and the Gaussian white noise data.

Table 4.2 The BDS test (W statistic) for the ASE returns, filtered returns, scrambled filtered returns and a Gaussian random surrogate.

		Original Data	**Filtered data (AR2 resid.)**	**Scrambled Residuals**	**Random Data**
Length (In Std.Dev.)	**Embedding Dimension**	**W Statistic (BDS/SD)**	**W Statistic (BDS/SD)**	**W Statistic (BDS/SD)**	**W Statistic (BDS/SD)**
2	2	27.04	25.004	1.2051	0.017541
2	3	28.706	27.415	0.90714	-0.85174
2	4	29.591	28.246	0.81774	-0.5836
2	5	30.149	28.891	0.50717	-0.30298
1.5	2	29.545	26.207	1.1245	0.27965
1.5	3	32.676	30.42	0.55202	-0.82804
1.5	4	34.562	32.42	0.082852	-0.57803
1.5	5	36.172	34.085	-0.3331	-0.18899
1	2	34.269	29.018	0.44556	0.39544
1	3	39.924	35.313	-0.21339	-0.86318
1	4	44.207	39.821	-0.64939	-0.59785
1	5	49.095	44.346	-0.75627	-0.10122
0.5	2	38.96	33.062	0.59975	0.88358
0.5	3	52.804	45.686	0.92414	-0.67752
0.5	4	68.825	59.08	0.7424	-0.40074
0.5	5	91.231	76.825	0.60395	0.15425

The corresponding results for the LSE series are presented in Table 4.3. In this case, too, the W statistic is significant at the 1% level for all different parameters used. However, its magnitude for both the original as well as the filtered LSE series is a little smaller compared to the ASE series.

Table 4.3 The BDS test (W statistic) for the LSE returns filtered returns and scrambled filtered returns.

		Original Data	**Filtered data (AR2 resid.)**	**Scrambled Residuals**
Length (In Std.Dev.)	**Embedding Dimension**	**W Statistic (BDS/SD)**	**W Statistic (BDS/SD)**	**W Statistic (BDS/SD)**
2	2	27.192	25.638	-0.38826
2	3	31.828	30.564	-0.8354
2	4	33.723	32.605	-1.2907
2	5	34.84	33.917	-1.5021
1.5	2	24.545	22.301	0.34898
1.5	3	28.84	26.807	-0.074984
1.5	4	31.033	29.124	-0.62213
1.5	5	32.69	31.002	-1.0509
1	2	21.358	18.533	0.68188
1	3	25.711	22.837	0.3756
1	4	28.458	25.577	-0.10826
1	5	31.082	28.348	-0.60564
0.5	2	18.488	15.627	0.78193
0.5	3	23.244	20.201	0.49116
0.5	4	26.621	23.48	0.19311
0.5	5	30.509	27.34	-0.29358

The results in Tables 4.2 and 4.3 indicate the existence of nonlinearities in the ASE and LSE daily returns. They are consistent with other studies that have applied the BDS test to financial data, e.g. to exchange rate data [Hsieh (1989)], to NYSE weekly stock returns [Scheinkman and LeBaron (1989)], to 90-day US Treasury Bill futures [Praschnik (1991)], to US daily returns of futures contracts [Yang and Brorsen (1993)] and to S&P 500 futures misprising series [Vaidyanathan and Krehbiel (1993)].

Recall that although nonlinearity is consistent with chaotic specifications BDS is not a test for chaos. The results above indicate only the existence of a non-iid and nonlinear structure in the return series of both markets, which can be of a chaotic or of a stochastic nature (e.g. ARCH or GARCH type processes).

Alternatively, we could say that the results of the BDS test suggest that a chaotic explanation cannot be ruled out for either of our series, so we can proceed with the rest of our testing framework to obtain a further insight of their underlying structure.

4.2.3 The Nonstationarity issue

In Econometrics, stationarity is related to the constancy of the moments of the unconditional distribution of a time series. In Financial Economics, stationarity is synonymous with structural changes in the markets due to the introduction of technological and financial innovations, policy changes etc. [Brock and DeLima (1995)]. In financial data, structural changes are possible, especially for long time periods, and there are studies supporting [Hsu, Miller and Wichern (1974), Pagan and Schwert (1990), Loretan and Phillips (1994)] or rejecting [Hinish and Patterson (1985)] the hypothesis that stock returns are nonstationary.

Given the evidence of non-IIDness and nonlinearity in our data by the use of the BDS test, it is important to determine whether these findings are due to structural changes. To do that, we followed a subsample analysis approach, also met in Hsieh (1991) and DeLima (1994).
Our ASE and LSE series have been divided into 3 and 6 approximately equal subsamples of 1000-1100 observations each, correspondingly. Assuming infrequent structural changes and using these shorter time periods, the effects of any structural changes should be removed or the period responsible for the nonlinear behavior should be isolated. Notice that the subperiod length of approximately 1000 observations, is a trade-off between short intervals and accurate BDS statistics.

The results for the ASE data are reported in Table 4.4 below. The best-fit linear model, selected by the same criteria described in the previous section has filtered each subsample.

As we can see in Table 4.4, each subperiod exhibits a different degree of linear dependence as the order of the AR model indicates. However, all subperiods have high W statistic values, significant at the 1% level, indicating that nonlinearity is strong and present throughout the series.

Table 4.4 The BDS test (W statistic) for 3 consecutive subperiods of the ASE daily returns.

		1st subsample [AR(1) Resid.]	*2nd subsample [AR(5) Resid.]*	*3rd subsample [AR(2) Resid.]*
Length (In Std. Dev.)	**Embedding Dimension (m)**	**W Statistic (BDS/SD)**	**W Statistic (BDS/SD)**	**W Statistic (BDS/SD)**
2	2	9.9967	14.323	12.3
2	3	11.948	16.214	13.579
2	4	12.891	16.545	14.219
2	5	13.035	17.19	14.525
1.5	2	11.198	15.801	12.787
1.5	3	12.214	17.507	14.778
1.5	4	13.004	18.337	16.076
1.5	5	13.412	19.32	17.133
1	2	11.198	17.065	13.026
1	3	12.214	19.186	16.518
1	4	13.004	20.381	19.197
1	5	13.412	22.024	22.001
0.5	2	12.319	17.565	13.579
0.5	3	13.233	21.444	19.662
0.5	4	13.983	24.665	25.858
0.5	5	14.331	29.383	33.985

The corresponding results concerning the subperiods in the LSE series are shown in Table 4.5. This time, the degree of linear dependence is the same across the different subperiods.

Nevertheless, the values of the W statistic compared to those for the ASE subperiods are in general lower, indicating a weaker presence of nonlinear dependence. For the 3rd period particularly, significance of the W statistic is questionable. Yet, despite the fact that differences in terms of the BDS statistic are more pronounced among the subperiods of the LSE series, the overall picture does not change and nonstationarity is not supported in these series, either.

Table 4.5 The BDS test (W statistic) for 6 consecutive subperiods of the LSE daily returns.

		1st subsample [AR(1) Resid.]	*2nd subsample [AR(1) Resid.]*	*3rd subsample [AR(1) Resid.]*	*4th subsample [AR(1) Resid.]*	*5th subsample [AR(1) Resid.]*	*6th subsample [AR(1) Resid.]*
Length (In Std.)	**Embedding Dimension**	**W Statistic (BDS/SD)**	**W Statistic (BDS/SD)**	**W Statistic (BDS/SD)**	**W Statistic (BDS/SD)**	**W Statistic (BDS/SD)**	**W Statistic (BDS/SD)**
2	2	6.7534	11.645	1.6473*	6.1493	11.864	2.9842
2	3	7.9987	13.656	2.8203	7.4577	13.717	4.2423
2	4	8.39	14.686	3.6267	7.8952	13.888	4.9798
2	5	8.5582	15.643	3.9515	8.4495	13.68	5.3454
1.5	2	5.8224	10.489	1.281*	5.8173	9.2927	1.9045*
1.5	3	6.9837	12.872	2.2396	7.1248	10.39	3.1831
1.5	4	7.3504	14.337	3.2055	7.6723	10.513	4.1626
1.5	5	7.8214	15.811	3.5987	8.3679	10.721	4.6679
1	2	5.0564	9.7364	0.67785*	5.2389	7.1547	1.2537*
1	3	6.595	12.598	1.5748*	6.8909	8.2853	2.4471
1	4	7.102	14.666	2.6479	7.8503	8.5804	3.5461
1	5	7.8467	17.119	3.1715	8.8864	9.3129	4.1944
0.5	2	4.3407	9.315	0.6614*	4.9322	6.3593	1.1039*
0.5	3	6.5456	13.058	1.2378*	6.9779	7.2168	2.2121
0.5	4	7.3922	16.557	1.6947*	8.347	7.1223	4.104
0.5	5	8.7286	21.026	1.7546*	10.395	7.9304	4.9376

* Not significant values at 5% level.

In all, the BDS test results support the non iid-ness and nonlinearity in both series (more pronounced in the ASE returns), which is not due to nonstationarity.

4.3 THE R/S ANALYSIS

4.3.1 Description of the method

Rescaled range analysis (R/S) is a very useful tool in investigating the fractal structure and non-periodic cycles in time series. Fractal structure is symptomatic of nonlinear feedback effects, resulting in repetitive structure in smaller and smaller scales. Stability is banded and the presence of long-term correlations implies structure. The type of hidden structure detected from R/S analysis is consistent with chaotic processes.

Technically, the R/S method includes the time dimension (standard econometrics apply to time series that are invariant with respect to time) by examining whether the range of fluctuations change, depending on the length of time used for measurement. Its origins are related to the «T to the one-half rule», that is, to the formula describing the Brownian motion (B.M.) :

$$R = T^{0.5} \tag{4.7}$$

Where: R = the distance covered by a random particle suspended in a fluid, and T = a time index

Equation (4.7) is commonly used in Finance to annualise volatility of e.g. monthly returns by multiplying monthly std. by $12^{0.5}$. It is also obvious that (4.7) shows how R is scaling with time T in the case of a random system and this scaling is given by the slope of the log(R) vs. log(T) plot, which is equal to 0.5. Yet, when a system or a time series is not independent, i.e. not a random B.M., (4.7) cannot be applied, so, Hurst gave the following generalisation which can be used in this case :

$$(R/S)_n = c\, n^H \tag{4.8}$$

where, $(R/S)_n$ = the Rescaled Range statistic measured over a time index n, c = a constant and H = the Hurst Exponent which shows how the R/S statistic is scaling with time.

The objective of the R/S method is to estimate the Hurst exponent, which, as we shall see, can characterise a time series. This can be done by transforming (4.8) to:

$$\log (R/S)_n = \log(c) + H \log(n) \tag{4.9}$$

and H can be estimated as the slope of the log/log plot of $(R/S)_n$ vs. n.

Given a time series $\{X_t : t=1,\ldots,N\}$, the R/S statistic can be defined as the range of cumulative deviations from the mean of the series, rescaled by the standard deviation. The analytical procedure to estimate the $(R/S)_n$ values, as well as, the Hurst exponent by applying (4.9), is described in the following steps :

Step 1: The time period spanned by the time series of length N, is divided into p contiguous subperiods of length n such that pn = N. The elements in each subperiod $X_{i,j}$, have two subscripts, the first (i=1,..,n) to denote the number of elements in each subperiod and the second (j=1,...,p) to denote the subperiod index. For each subperiod j the R/S statistic is calculated, as:

$$\left(\frac{R}{S}\right)_j = s_j^{-1}\left[\max_{1\le k\le n}\sum_{i=1}^{k}(X_{ij}-\overline{X}_j) - \min_{1\le k\le n}\sum_{i=1}^{k}(X_{ij}-\overline{X}_j)\right] \qquad (4.10)$$

where s_j is the standard deviation for each subperiod.

In (4.10), the k deviations from the subperiod mean have zero mean, hence the last value of the cumulative deviations for each subperiod will always be zero. Due to this, the maximum value of the cumulative deviations will always be greater or equal to zero, while the minimum value will always be less or equal to zero. Hence the range value [the bracketed term in (4.10)], will always be nonnegative.

Normalizing (rescaling) the range is important since it permits diverse phenomena and time periods to be compared, which means that R/S analysis can describe time series with no characteristic scale.

Step 2: The $(R/S)_n$, i.e. the R/S statistic for time length n, is given by the average of the $(R/S)_j$ values for all the p contiguous subperiods with length n, as:

$$\left(\frac{R}{S}\right)_n = \frac{1}{p}\sum_{j=1}^{p}\left(\frac{R}{S}\right)_j \qquad (4.11)$$

Step 3: Equation (4.11) gives the R/S value which corresponds to a certain time interval of length n. In order to apply equation (4.9), steps 1 and 2 are repeated by increasing n to the next integer value, until n = N/2, since, at least two subperiods are needed, to avoid bias.

From the above procedure, it becomes obvious that the time dimension is included in the R/S analysis by examining whether the range of the cumulative deviations depends on the length of time used for the measurement. Once (4.11) is evaluated for different n periods, the Hurst exponent can be estimated through an ordinary least squares regression from (4.9).

The Hurst exponent takes values from 0 to 1 ($0 \leq H \leq 1$). Gaussian random walks, or, more generally, independent processes, give H = 0.5.

If $0.5 \leq H \leq 1$, positive dependence is indicated, and the series is called persistent or trend reinforcing and in terms of equation (4.7) the system covers more distance than a random one. In this case, the series is characterised by a long memory process with no characteristic time scale. The lack of a characteristic time scale (scale invariance) and the existence of a power law (the log/log plot) are the key characteristics of a fractal series.

If $0 \leq H \leq 0.5$, negative dependence is indicated, yielding anti-persistent or mean-reverting behaviour (only if the system under study is assumed to have a stable mean). In terms of equation (4.7), the system covers less distance than a random series, which means that it reverses itself more frequently than a random process.

A Hurst exponent different from 0.5 may characterise a series as fractal. However a fractal series might be the output of different kinds of systems. A «pure» Hurst process is a fractional Brownian motion [Mandelbrot, (1972), Mandelbrot and Wallis (1969a,d,c,d)] also known as biased random walk, or fractal noise, or coloured (black) noise, that is, a random series the bias of which can change abruptly but randomly in direction or magnitude.

However, chaotic systems have also Hurst exponents H>0.5, and in chaotic terms long memory effects correspond to sensitive dependence on initial conditions. As already mentioned, the SDIC property combined with fractality characterises chaotic systems. Pure chaotic processes have Hurst exponents close to 1.

When dealing with real data the problem of distinguishing between the above mentioned alternatives becomes more difficult due to the existence of noise.

Most series are contaminated by either additive or dynamical (system) noise. Hence, in most cases, which include financial data, the problem is to distinguish between fractal noise and noisy chaos.

R/S analysis provides a very useful tool to solve the above problem since it is extremely robust to both additive and system noise [Peters (1994)]. Noise lowers the H value of a series and obscures the difference e.g. between a fractional Brownian motion with H = 0.70 and a chaotic process which originally has H=0.92 but noise contamination reduces it to 0.70, as well. However, R/S analysis is able to detect, even when noise is present, the existence of cycles in the series and thus to characterise it.

Cycles can be either periodic or non-periodic in the sense that the system has no absolute frequency. Non-periodic cycles can be further divided to statistical cycles and chaotic cycles. Fractal noises exhibit statistical cycles, i.e. cycles with no average cycle length. Actually, they are random cycles of different length due to long-run correlations and randomly changing bias[2] of the system. On the contrary, deterministic systems in the form of chaotic flows, such as the Mackey-Glass equation[3], or noisy chaotic processes have chaotic cycles, which have an average frequency. In this case the cycle denotes the finite memory length of the system, which measures the folding of the attractor.

In general, fractal noises will have no discernible cycles but in practice and in a certain time scale Fractional Brownian Motion (FBM) might exhibit a finite memory effect which is usually a statistical artifact due to the limited length of the series examined. In this case, fractal noise can be distinguished from a chaotic alternative by examining whether the cycle is independent of the time scale used. Cycles that do not depend on the sample size (are independent of the time scale) indicate the noisy chaos alternative.

---

2 Abrupt changes in direction due to exogenous events, predictable or not.

3 Mackey and Glass (1977) have produced the following delay differential equation: $dX/dt=aX(t-r)/[1+X(t-r)^c]-bX(t)$. The state space of this system is infinite-dimensional but its attractor has a finite dimension.

There are alternative ways to detect cycles by the use of R/S analysis and estimate their length:

- Once the data is long enough, the cycle can be discerned from the log(R/S) vs. log(n) plot as a "break-point". At this point, if a cycle exists, the slope of the curve in the log(R/S) vs. log(n) plot should change and the plot crosses to a random walk ($H \cong 0.5$). Since log(n) measures time on the X-axis, the length of the cycle can also be measured by the X-coordinate of the break point.
- Alternatively, the regressions of log(R/S) vs. log(n) [see (4.9)] can be used for different n-lengths. The n value, for which a peak value of H is obtained, corresponds to the cycle length.
- Finally, a third and more efficient way to find the cycle length is given by the V-statistic [Hurst (1951), Peters (1994)] defined as:

$$V_n = (R/S)_n / n^{0.5} \quad , \qquad (4.12)$$

The V_n vs. log(n) plot gives a flat line for an independent random process and an upwardly sloping curve in the case of persistent series. The existence of a cycle and its length can be discerned (even when noise is present) from the "break-point" in this plot occurring when V_n reaches a peak and then flattens out, an indication that the long-memory process has dissipated.

4.3.2 Application problems

The application of R/S analysis is related to specific problems, which should be discussed.

The first problem is related to the procedure of step 1 described above. In order to create the p subperiods, the series of N-length should be divided by n (the length of each subperiod) as:

$$p = N/n \qquad (4.13)$$

In (4.13), p should always be an integer since all the subperiods must have the same length n. Since N is fixed and n is changing, some points should be left out when (4.13) gives no integer p. Obviously, the number of these "left-out" points, depends on the length of the data N and the divisor n, and in some cases exceeds 30% or more of the total points in the data, resulting in significant bias and unreliable H estimates. To account for this, Peters (1992,1994) suggests that only the time increments that include both the beginning and the ending points of the data should be used, i.e. only the n values that produce integer p values. However, these n values might be very few, especially when the N-length is small (N<5000) which means that there are not enough regressors to estimate H in (4.9). Hence, for small data sets there is a "trade-off" between adequate number of regressors and bias induced by "left-out" points. To account for this, the N-length is adjusted in each case to give the maximum number of perfect divisors and a very small number of "left-out" points, in order to minimize bias.

A second problem is related to the **evaluation of the H exponent from a statistical point of view**. Specifically, we should be able to asses whether an H value is statistically significant in comparison to a random null, i.e. to the H exponent exhibited by an independent random system. Peters (1994) shows that under the Gaussian null, a modification of a formula developed by Anis and Lloyd (1976) allows for hypothesis testing by computing $E(R/S)_n$ and E(H), the expected variance of which will depend only on the total sample size N, as Var (H) = 1/N.

However, if the null is still iid randomness (strict white noise) but not Gaussianity, formal hypothesis testing is not possible. This problem has been brought up by Peters, but it has not been solved, thus leaving room for disputing R/S analysis results in the existing literature. To overcome this problem we used bootstrapping (and followed the process described in Chapter 3) to assess the statistical significance of the H exponents of our series, against both the Gaussian and the iid random null hypotheses.

To test against the Gaussian random null, the H exponent from 5000 random shuffles of a Gaussian random surrogate, having the same length, mean and variance with our return series, is calculated and compared to the test statistic, i.e. the actual H exponent of our series.

If the latter is found to be greater than 0.5 and persistence of the series is possible, then the null hypothesis tested is formed as: $H_0 : H = H_G$ (and the alternative as H_1: $H > H_G$) which can be expressed as *"The actual H estimate from the series tested is equal to (is not significantly different from) the* H_G *estimate from a Gaussian random data with* *the same length, mean and std as the series tested".*

The significance level of the test is constructed as the frequency with which the test statistic (the H estimate) from the Gaussian surrogates is greater than or equal to the actual statistic for the original data [Noreen (1989)]. The null hypothesis is rejected if the significance level is smaller than the conventional rejection levels of 1%, 2,5% or 5%.

To test against the iid null, the same procedure is followed but this time we randomize the series tested to produce 5000 iid random samples having the same length and distributional characteristics as the original series. In this case, rejection of the null means that the actual H exponent calculated from the original series is significantly greater than the one calculated from an iid random but non-Gaussian series. Hence, this is also a test for non-iid-ness.

A third and more serious problem is related to the **sensitivity of R/S analysis to short-term dependence**, which can also lead to unreliable results [Wallis and Matalas (1970), Milonas et.al. (1985), Davies and Harte (1987), Aydogan and Booth (1988), Lo and Mackinlay (1988), Lo (1991), Haubrich and Lo (1989)]. Peters (1994), shows that Autoregressive (AR), Moving Average (MA) and mixed ARMA processes exhibit Hurst effects, but once short-term memory is filtered out by an AR(1) specification, these effects cease to exist. On the contrary, ARCH and GARCH models do not exhibit long term memory and persistence effects at all. Hence, a series should be pre-filtered for short-term linear dependence prior to applying the R/S analysis.

In our analysis, as already discussed in the case of the BDS test, we use partial autocorrelograms and Schwarz's (1978) information criterion to indicate the linear AR filter of our data.

An alternative way to account for short-term dependence is to use the R/S test statistic modified by Lo (1991).

In Lo's modification, short-range dependence is incorporated into the partial sum variance estimator to the denominator of the classical Mandelbrot's (1972) R/S statistic as:

$$Q_q = \hat{s}_N^{-1}(q)\left[\max_{1\le k\le N}\sum_{t=1}^{k}(X_t - \overline{X}_N) - \min_{1\le k\le N}\sum_{t=1}^{k}(X_t - \overline{X}_N)\right] \qquad (4.14)$$

where

$$\hat{s}_N^2(q) = \hat{s}_x^2 + 2\sum_{t=1}^{q} w(q)\hat{g}_t, \qquad w_t(q) = 1 - \frac{t}{q+1}, \quad q<N \qquad (4.15)$$

and $\hat{s}_x^2$ and $\hat{g}_t$ being the variance and autocovariance estimators of X.

According to (4.14) and (4.15), if $\{X_t\}$ is subject to short-range dependence, the estimator $\hat{s}_N(q)$ involves the sums of squared deviations of X and its weighted autocovariances up to lag q. This latter term provides the modification of the original $(R/S)_n$ statistic.

This test, unlike classical R/S analysis described above, does not have to rely on subsample analysis. The test's null is short term dependence, which operationally is defined by Lo as a "strong-mixing" process, a notion due to Rosenblatt (1956),[4] in order to derive the asymptotic distribution of Q_q.

Lo shows that under certain conditions, which place restrictions on the maximal moments, the degree of distributional heterogeneity and the maximal degree of dependence in $\{X_t\}$, the statistic $V_q = N^{-(1/2)} Q_q$ converges to the range of a "Brownian bridge" on the unit interval, a well defined random variable with mean $(\pi/2)^{(1/2)}$, variance $\pi^2/6-\pi/2$ and a positively skewed distribution function.

The critical values of the test derived by the asymptotic cumulative distribution function are given in Table 4.6.

Table 4.6 Asymptotic critical values of the modified R/S statistic

Probability level	0.5%	2.5%	5%	10%	90%	95%	97.5%	99.5%
Critical value	0.721	0.809	0.861	0.927	1.620	1.747	1.862	2.098

The main advantage of the test is that it allows for formal statistical testing and is robust against serial correlation and some forms of nonstationarity. It is specifically designed to distinguish between weakly dependent processes, e.g. ARMA, and strongly dependent processes, e.g. fractionally integrated (ARFIMA) models [Hosking (1981), Granger and Joyeux (1980), Cheung (1993)]. Notice that the main characteristic of these strongly dependent processes is the slowly decaying autocorrelation functions and non-periodic cyclical patterns.

Moreover, the test's null is wide enough to include models with longer-term correlations such as the stochastic models of persistence proposed by Campbell and Mankiw (1987), Fama and French (1988) and Poterba and Summers (1988).

The main disadvantage of the test is that, unlike the classic R/S analysis, it is not able to specify the cycle length of the series tested. In addition, there are certain shortcomings related to the test "per se". Lo (1991) shows that there are forms of short-term dependence violating the assumptions of the test's null[5]. He also reports low power of the test against chaotic processes such as the "tent-map", a long-range dependent process with very low autocorrelation.

Hiemstra and Jones (1995) find that right and left-tailed bootstrapped critical values of the modified R/S statistic fall below their asymptotic counterparts, resulting to higher right-tailed and lower left-tailed rejection rates. According to their analysis, this is due to the test sensitivity to moment condition failure, i.e. to the magnitude of the maximal moment of their series, which is less than 4.

Brock and DeLima (1995) use Monte Carlo simulation to also find that the sampling distribution of the test is shifted to the left relatively to the asymptotic distribution.

Another problem is related to the sensitivity of the test to the truncation lag-parameter q in (4.15). Lo (1991) employs Monte Carlo simulations to assess the power of the test, which declines with increasing q and decreasing sample size. In fact, even for sample sizes of N=1000, the empirical rejection rates were much lower than nominal sizes for q values exceeding $(N)^{1/3}$.

---

4 Strong mixing requires that the maximal dependence between two events becomes trivially small as their separation time increases without bound.

5 For example, the test has no power against processes with maximal moments less than 4 violating the moment condition of the test, or the first difference of a stationary process violating the heterogeneity condition.

Little is known about the optimal choice of q, although Andrews (1991) suggests a data dependent formula given by:

$$q_N = 1 + \mathrm{INT}\left[\left(\frac{3N}{2}\right)^{\frac{1}{3}} \cdot \left(\frac{2\hat{r}}{1-\hat{r}}\right)^{\frac{2}{3}}\right] \tag{4.16}$$

where r = first order autocorrelation coefficient

However, the truncation lag given by this formula is optimal only for an AR(1) data generating process.

Although R/S analysis, combined with bootstrapping for assessing the statistical significance of the H exponent, provides a powerful tool for detecting persistent behavior and long-term cycles, we also employed the modified R/S statistic to cross-check our findings.

In our application of the modified R/S statistic, we used different q-lengths equal to q_N=INT$[(N)^{1/4}]$, INT$[(N)^{1/3}]$, INT$[(N)^{1/2}]$, 100, 150, as well as the q values derived from the data-dependent formula in (4.16).

In addition, to assess the test's results, we have used the asymptotic and bootstrapped critical values of our series. The latter are based on the test statistics derived from 5000 time-scrambled shufflings of our data, producing iid series consistent with the test's null and robust to violations of the moment condition of the test.

The empirical literature of R/S analysis includes different studies with contradictory results. Classical R/S analysis has been applied to employment series [Booth and Koveos (1983)], gold market prices [Booth et.al. (1982)], exchange rates [Booth et. al. (1980, 1982b), Peters (1991a)], stock market indices [Peters (1989, 1991a,b, 1992,1994)] and common stock returns [Green and Fielitz (1977)]. In all these cases long-term memory has been reported.

On the other hand, the modified R/S statistic has been applied to macroeconomic data [Haubrish and Lo (1989), Mills (1992)], gold market prices [Cheung and Lai (1993)], currency futures markets [Kao and Ma (1992)], stock indices [Lo (1991), Mills (1993), Cheung et. al. (1993), Crato (1994)] and common stock returns [Hiemstra and Jones (1995)], and weak or absence of long-term dependence has been reported by all of these studies. Since both approaches have certain advantages and disadvantages, our empirical application incorporates both, as already discussed.

4.3.3 Empirical evidence

I.The ASE Return series

As already mentioned, the ASE returns consist of 3181 daily observations covering a 13-year period. Since the series exhibit serial autocorrelation the AR(2) filtered series have been also used in our analysis.

Figures 4.4a and 4.4b present the log-log plot of the R/S statistic versus time (n-days) and the semi-log plot of the V-statistic respectively for four different series. The original returns, the filtered returns, a random iid series of scrambled filtered returns and a random Gaussian version of the filtered returns, the last two being presented for visual comparison purposes.

The V-statistic plot shows a possible long-term cycle with a length of 1040 trading days for both the original and filtered series, while for the scrambled residuals and the Gaussian surrogate a much flatter curve occurs. However, a smaller cycle appears also, having a length of approximately 155 trading days. As we can see in Figure 4.4b, the V-statistic curve for the raw and filtered series rises up from point A to point B corresponding to 155 days, then from B to C (approx. 390 days) it flattens out to rise up again to point D (1040 days).
To verify these findings the Hurst exponent (H) has been calculated for each cycle, as well as for the intermediate - between the two cycles - period and bootstrapping has been used to assess the statistical significance of the H estimates against both the Gaussian random and the random iid alternatives.

The relevant results are presented in Table 4.7. In this Table, capital letters in parentheses beside each column number correspond to the time periods signified in the V-statistic plot. Part A of the Table 4.7 presents the R/S analysis findings. We can see that the H estimates from the filtered series for all the different periods tested are lower than the ones from the original data, due to the linear autocorrelation bias that has been removed.

R/S Analysis / ASE Daily returns

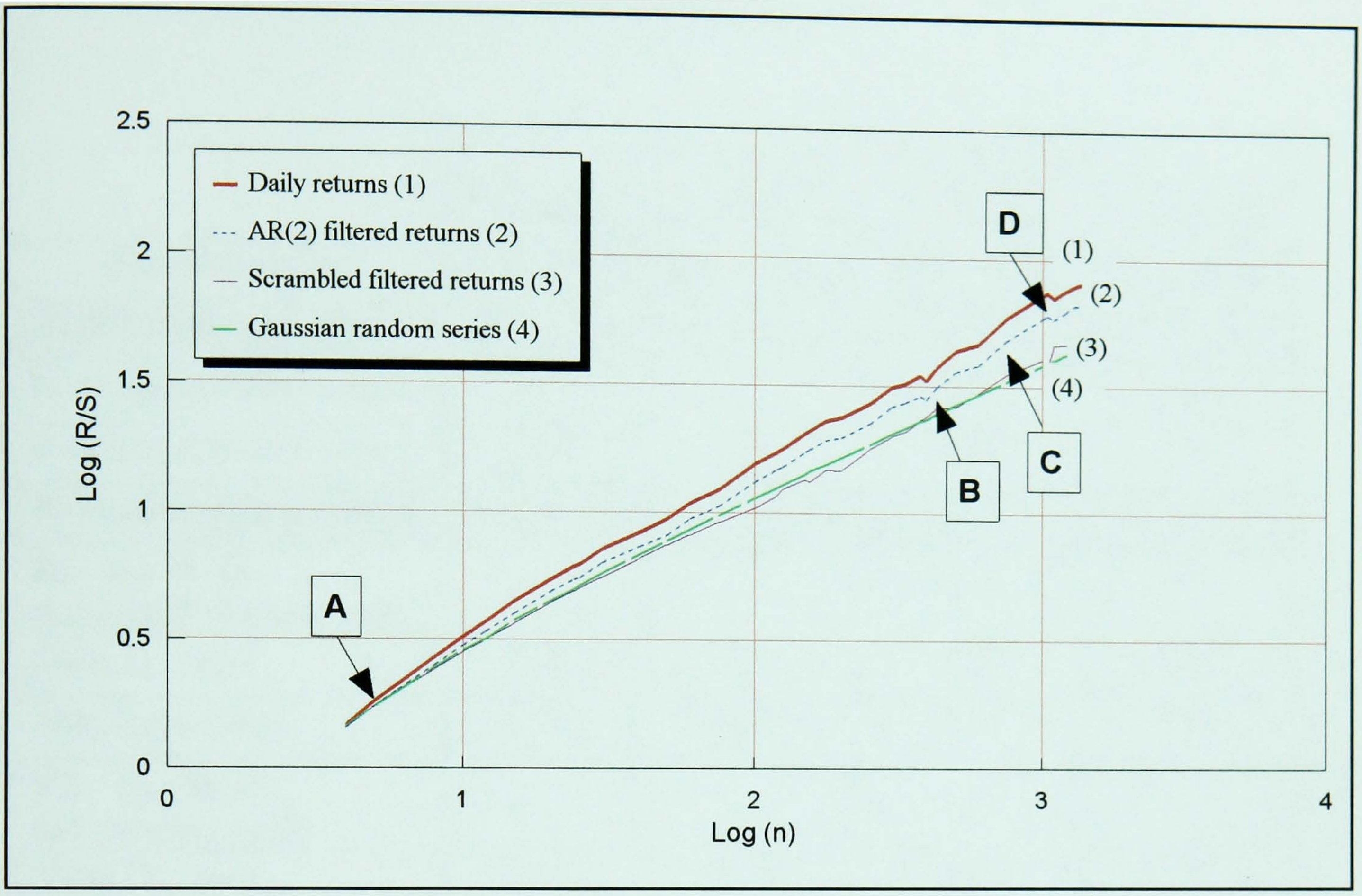

Figure 4.4a Log(R/S) vs. Log(n) plot for the H exponent estimation of the raw and filtered ASE returns, the scrambled filtered returns and a Gaussian surrogate.

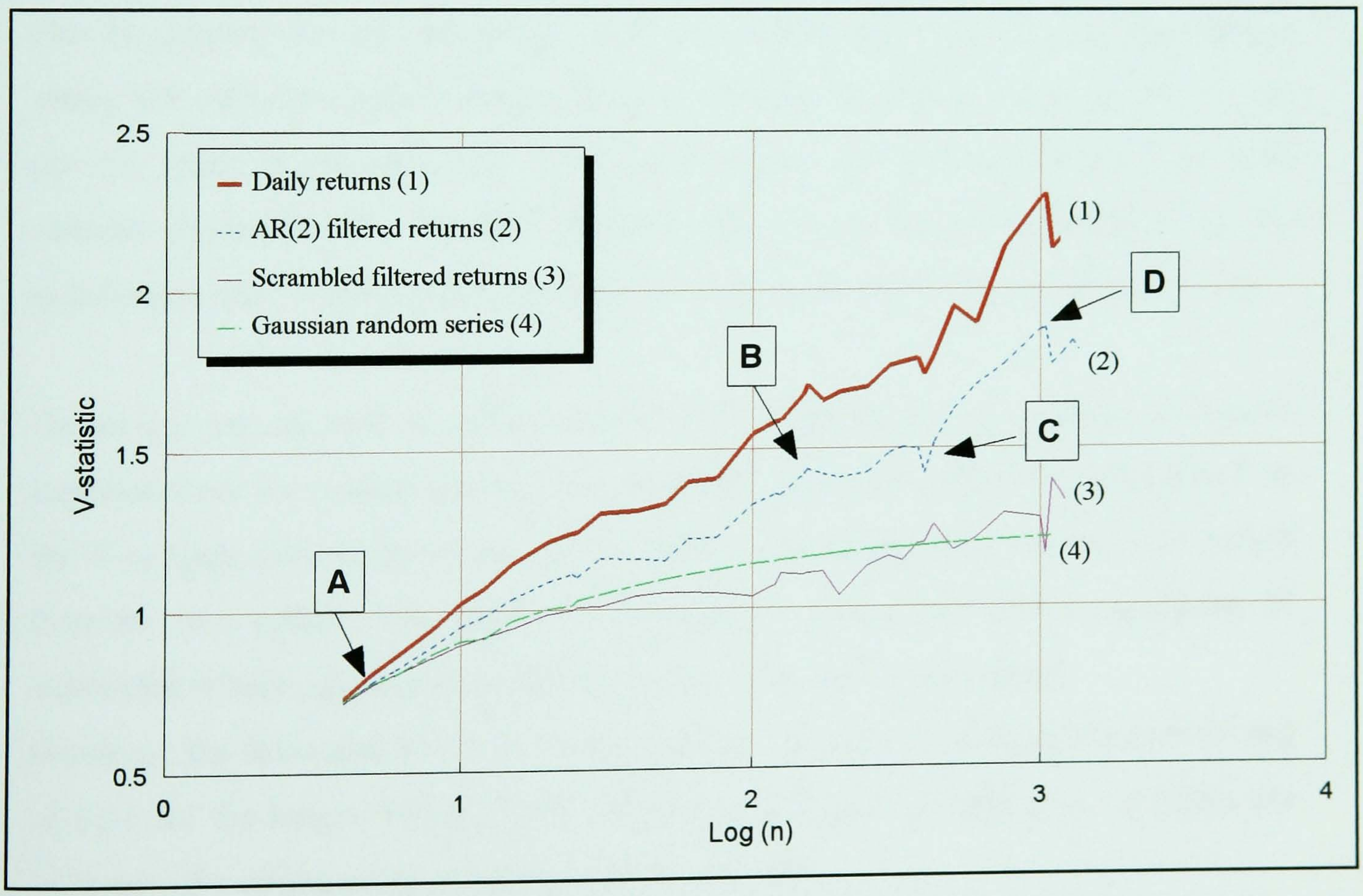

Figure 4.4b V-statistic vs. Log(n) plot for the cycle-length estimation of the raw and filtered ASE returns, the scrambled filtered returns and a Gaussian surrogate.

Table 4.7 Hurst estimates test of significance against two random alternatives of the daily ASE returns

	Time period (in trading days)			
	1 (A to B)	2 (B to C)	3 (C to D)	4 (A to D)
One-day returns	**1<n<155**	**156<n<393**	**394<n<1040**	**4<n<1040**
A. R/S analysis results				
H estimate(original series)	0.67	0.57	0.74	0.66
H estimate (filtered series)	0.64	0.55	0.73	0.63
B. Bootstrapping results				
B1. H_0: $H=H_G$ (Gaussian random null)				
Mean H_G value	0.60	0.53	0.52	0.56
Significance level	0.003	0.392	0.025	0.005
B2. H_0 : $H=H_R$ (iid random null)				
Mean H_R value	0.59	0.53	0.54	0.56
Significance level	0.001	0.415	0.018	0.003

Part B, presents the bootstrapping results of testing the H estimate from the filtered series, for each time period, against the two random alternatives. In the Table we can see the mean H estimate from 5000 bootstrapped samples corresponding to each random alternative (H_G for the Gaussian and H_R for the iid) and below it, the significance level which defines whether the random null hypothesis is rejected or not.

As we can see, the null, according to which the actual H estimate is equal to the one expected from the random alternatives, cannot be rejected only for the period B to C in the V-statistic plot, verifying that this is indeed a no-Hurst effect intermediate period between two cycles. The latter are no artifacts, since they exhibit significant H exponents at high significance levels (less than 1% in most of the cases).
However, the H estimates of 0,64 for the smaller 8-month (155 trading days) cycle and of 0,63 for the longer 4-year (1040 trading days) cycle, are quite low, reflecting the existence of a strong noisy component in the daily series.

The results from the modified R/S statistic seem to confirm the above findings concerning the fractality and long-term dependence of the ASE series.

Table 4.8, presents the estimates of the V_q statistic for different q values up to q = 150. We used both the asymptotic as well as the bootstrapped critical values to assess the statistical significance of the test statistic. The right-tail bootstrapped critical values (at the 1%, 2,5% and 5% significance levels), corresponding to each one of the q values, are also presented in the lower part of Table 4.8 These values are found to be lower than the asymptotic ones, confirming the findings of Hiemstra and Jones (1995)

Table 4.8 The modified R/S statistic of the ASE return series & bootstrapped critical values of the test statistic.

q	**Andrew's** 12	$N^{1/4}$ 8	$N^{1/3}$ 15	$N^{1/2}$ 56	**100**	**150**
V_q – statistic	2.165^{*}	2.244^{*}	2.122^{*}	1.793^{***}	1.765^{***}	1.749^{***}
Bootstrapped Critical Values						
Significance level / q	**12**	**8**	**15**	**56**	**100**	**150**
1.0%	1.918	1.928	1.917	1.845	1.820	1.819
2.5%	1.811	1.814	1.804	1.774	1.741	1.724
5.0%	1.711	1.721	1.682	1.664	1.658	1.651

* Significance at 1.0% level according to the asymptotic critical values

** Significance at 2.5% level according to the asymptotic critical values

*** Significance at 5.0% level according to the asymptotic critical values

We can see from Table 4.8 that for all the q values employed, the V_q estimates reject the test's short-range dependence null when compared to either the asymptotic or the bootstrapped critical values of the test. Specifically, when the asymptotic values are used, significance varies from 1% for the lower q values to 5% for the higher ones. However, when the bootstrapped values are used, significance does never exceed 2.5% for the whole q range.

The remaining step is to investigate whether the cycles that have been found are true nonperiodic cycles compatible with a noisy chaos explanation or just a stochastic boundary due to data size, compatible with fractional Brownian motion with finite memory. As already discussed, R/S analysis can make that distinction and a cycle independent of the sample size is a clear indication towards the first alternative.

In this respect, R/S analysis is performed with 5-day and 20-day returns produced from our daily data. To avoid autocorrelation bias both series were pre-filtered with an AR(2) and an AR(1) filter respectively, as indicated by partial autocorrelograms and Swartze's information criterion.

Figures 4.5a, 4.5b and 4.6a, 4.6b present the R/S log-log plot and the V-statistic plot for the 5-day and 20-day returns respectively. As can be clearly seen from these plots, both series exhibit cyclycity identical to the one observed in the daily data. Specifically, in the case of the 5-day series, 2 cycles of 31 (points A to B) and 208 (points C to D) 5-day intervals, and a flattening period between them (points B to C) is pictured in the V-statistic plot 4.4b.

In the case of the 20-day series, 2 cycles of 8 (points A to B) and 51 (points C to D) 20-day intervals, and a flattening period between them (points B to C) is pictured in the V-statistic plot 4.5b.

R/S Analysis / ASE Five - day returns

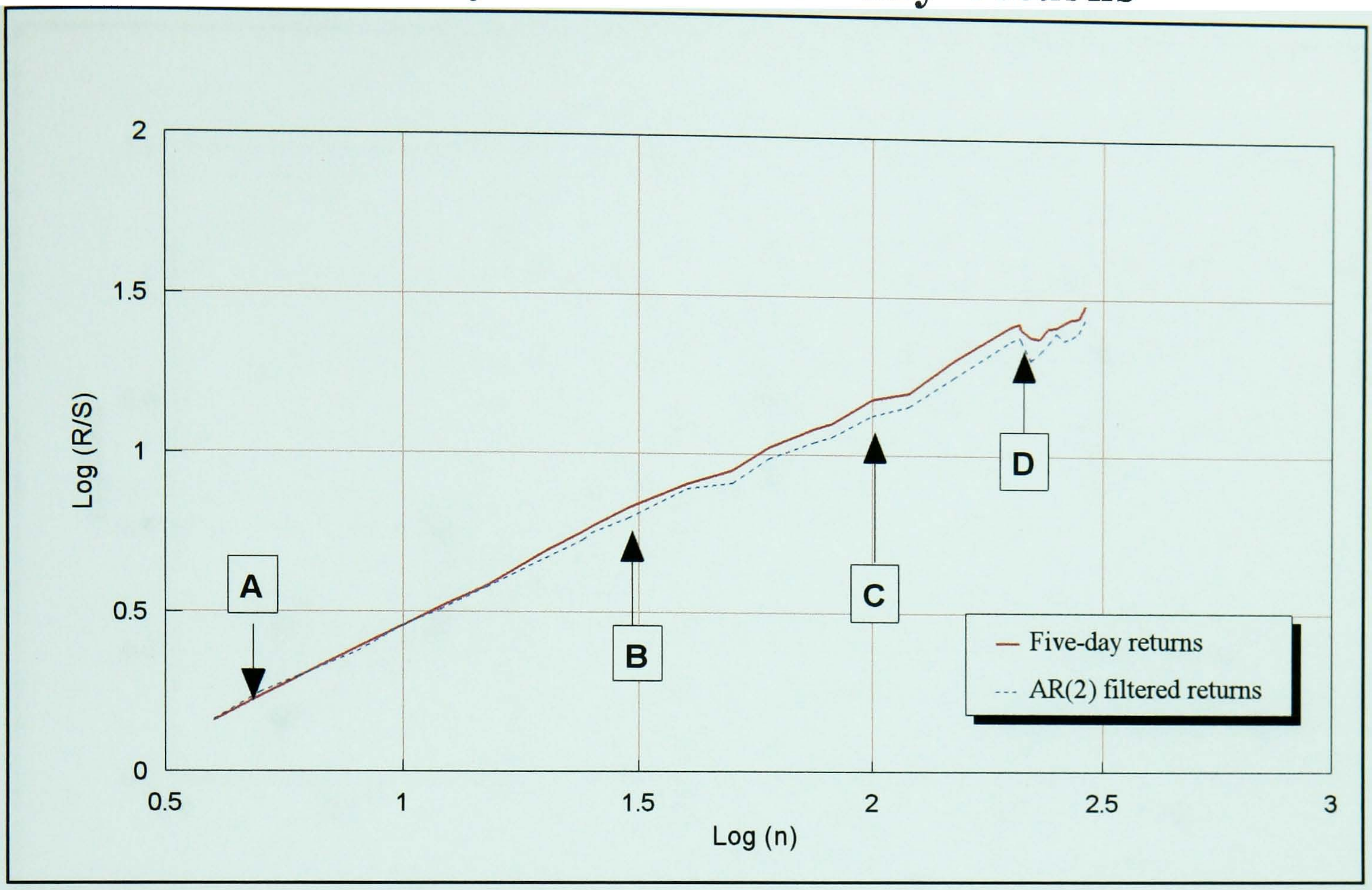

Figure 4.5a. Log(R/S) vs. Log(n) plot for the H exponent estimation of the raw and filtered ASE 5-day returns.

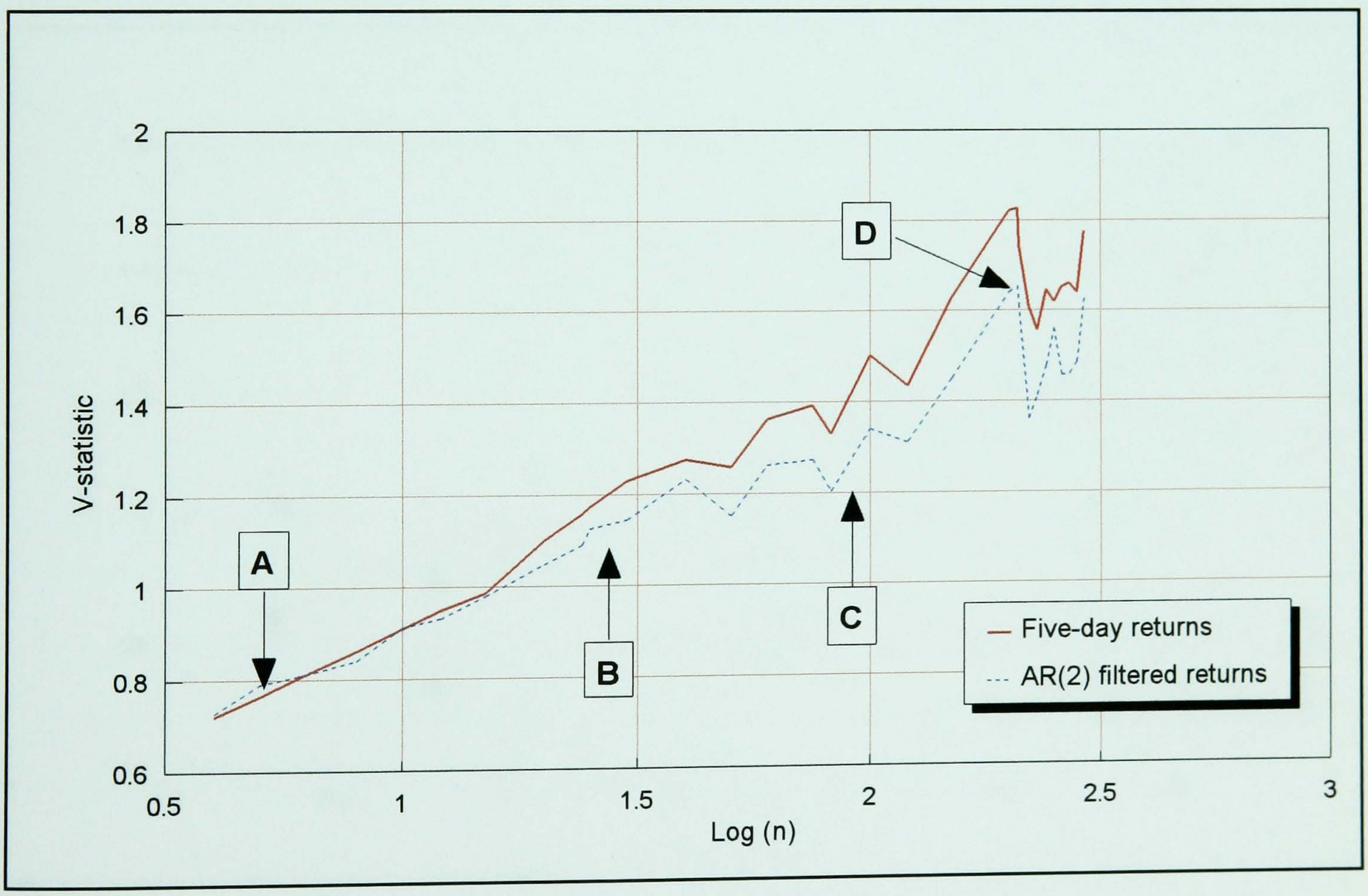

Figure 4.5b. V-statistic vs. Log(n) plot for the cycle-length estimation of the raw and filtered ASE 5-day returns.

R/S Analysis /ASE Twenty - day returns

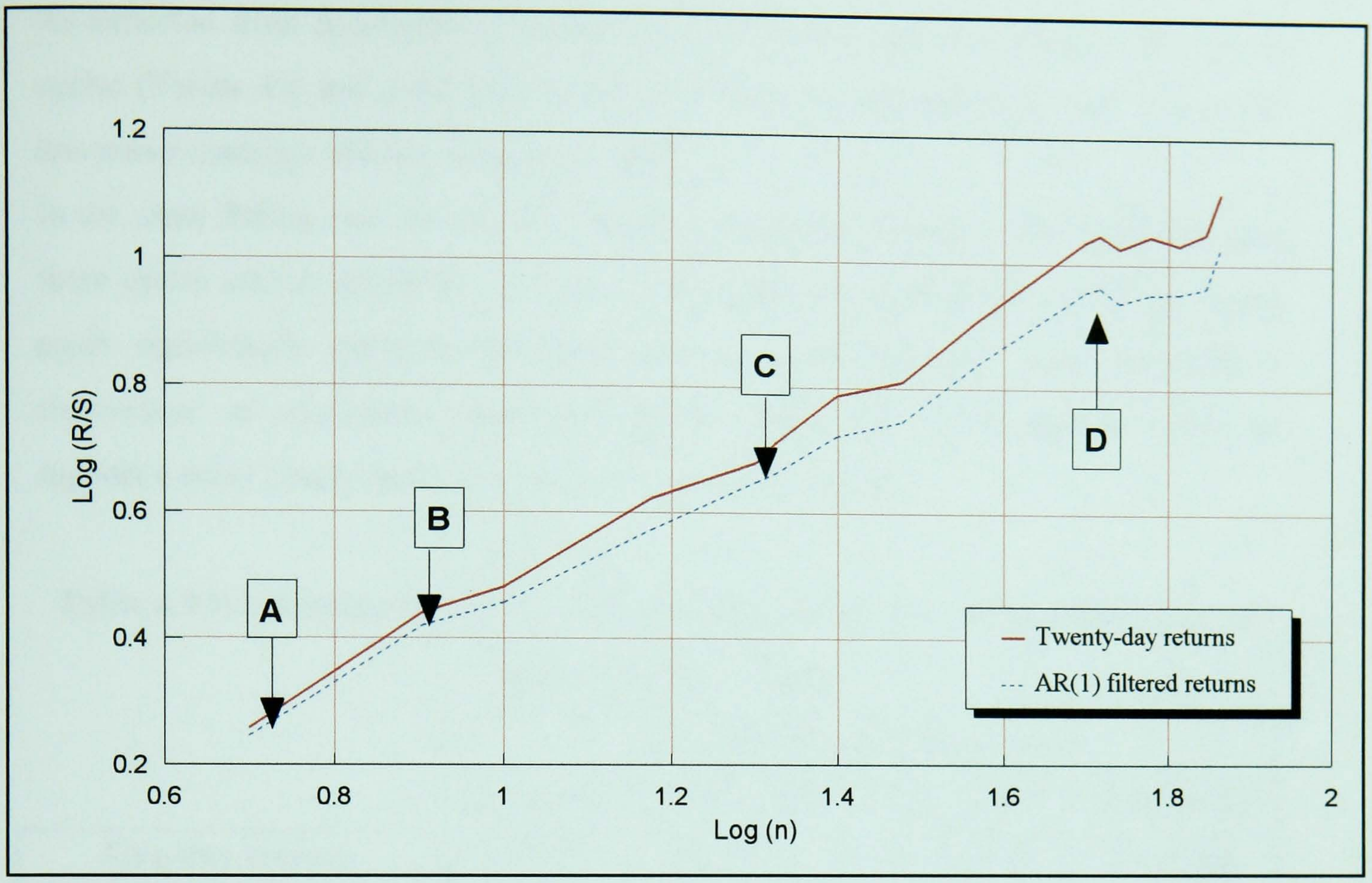

Figure 4.6a. Log(R/S) vs. Log(n) plot for the H exponent estimation of the raw and filtered ASE 20-day returns.

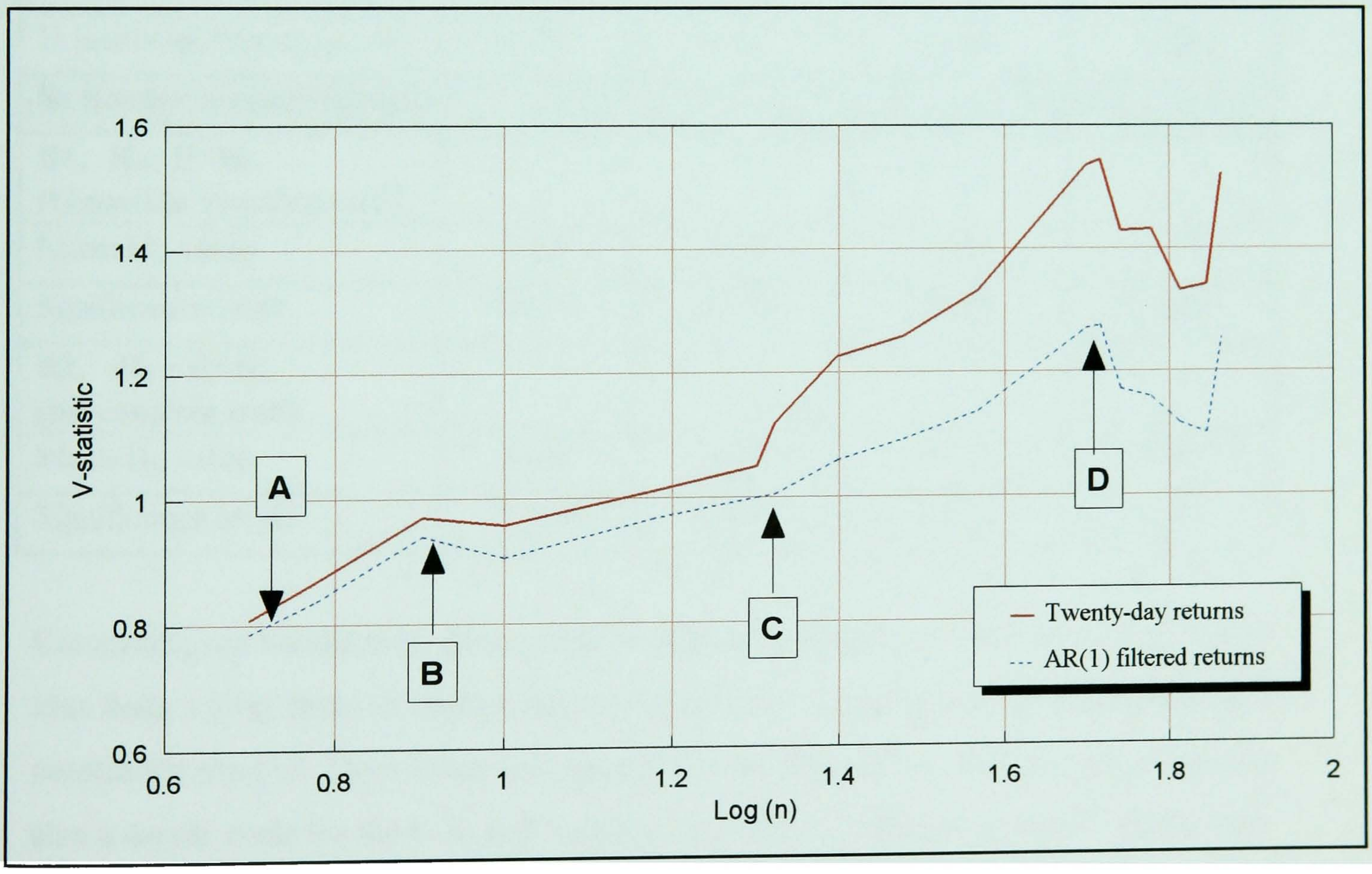

Figure 4.6b. V-statistic vs. Log(n) plot for the cycle-length estimation of the raw and filtered ASE 20-day returns.

As expected from non-random iid data, the H exponents corresponding to the above cycles (Tables 4.9 and 4.10) are increasing with longer sampling intervals, due to the less noisy character of lower frequency data.

In the same Tables, we can see that bootstrapping results support the hypothesis that these cycles are not artifacts, since the H exponents corresponding to them are found again significantly different (greater) than the expected ones from the random alternatives, at acceptable significance levels. Hence, the above analysis seems to support a noisy chaos alternative, against a fractal noise one.

Table 4.9 Hurst estimates and test of significance against two random alternatives of the five-day ASE returns

	Time period (in trading days)			
	1 (A to B)	**2 (B to C)**	**3 (C to D)**	**4 (A to D)**
Five-day returns	**1<n<31**	**32<n<82**	**83<n<208**	**1<n<208**
A. R/S analysis results				
H estimate(original series)	0.79	0.62	0.85	0.72
H estimate(filtered series)	0.72	0.57	0.80	0.68
B. Bootstrapping results				
B1. H_0: $H=H_G$ (Gaussian random null)				
Mean H_G value	0.66	0.57	0.56	0.57
Significance level	0.015	0.495	0.028	0.018
B2. H_0 : $H=H_R$ (iid random null)				
Mean H_R value	0.65	0.56	0.55	0.60
Significance level	0.008	0.444	0.024	0.016

Comparing our results with the existing literature, we should notice that Peters (1994), also finds a long cycle of approximately 4 years and a smaller one of approximately 2 months for the U.S. Dow Jones Industrials. In a previous study, Peters (1991a) reports also a 4-year cycle for the U.S. and Japan stock markets. Finding the same cycle for the U.S and Japan stock markets, which are mature and efficient markets compared to ASE which is an emerging one, might be interesting, since it reveals the possibility of similar structures between different markets. If this is true, a possible connection of the 4-year cycle to the cycles in the economy should also be considered.

However, notice that Peters uses a very long U.S. data set covering more than 30 years and claims that data should cover at least 10 cycles for safe conclusions. Thus, in our case the existence of the longer cycle should be viewed with caution, although the shorter one fulfills this criterion, as well.

Table 4.10 Hurst estimates and test of significance against two random alternatives of the twenty-day ASE returns

	Time period (in trading days)			
	1 (A to B)	**2 (B to C)**	**3 (C to D)**	**4 (A to D)**
Twenty-day returns	**1<n<8**	**9<n<20**	**21<n<51**	**1<n<51**
A. R/S analysis results				
H estimate(original series)	0.97	0.70	0.85	0.77
H estimate(filtered series)	0.91	0.62	0.78	0.72
B. Bootstrapping results				
B1. H_0: $H=H_G$ (Gaussian random null)				
Mean H_G value	0.77	0.65	0.57	0.60
Significance level	0.005	0.621	0.048	0.012
B2. H_0 : $H=H_R$ (iid random null)				
Mean H_R value	0.75	0.63	0.57	0.61
Significance level	0.001	0.536	0.055	0.014

In conclusion, fractality, persistence and long-term dependence are supported for the ASE return series, when a thorough R/S analysis is performed, enhanced by bootstrapping methodology to assess significance of the estimated Hurst exponents and by the modified R/S statistic to reject short-term dependence alternatives. Moreover, a noisy chaos alternative is favoured against a coloured noise one like Fractal Brownian Motion.

II.The LSE Return series

The LSE returns consist of 6650 daily observations covering a 25-year period. As already discussed, the series exhibit serial autocorrelation and an AR(1) filtered series has been used in our analysis. Figures 4.7a and 4.7b present the log-log plot of the R/S statistic versus time (N-days) and the semi-log plot of the V-statistic respectively, for the filtered series as well as for a Gaussian random version of this data.

The V-statistic plot does not show a clear picture regarding the possible cycle(s) of the data as it was the case of the Greek series. In fact, the V-statistic curve of the Gaussian series is quite close to the one of the filtered series casting doubt on the existence of significant difference between them.

However, two possible alternative[6] cycles can be discerned, the first having length of 330 trading days (points A to B in fig. 4.7b) and the second of 660 trading days (points A to C in fig. 4.7b).

As we did in the case of the Greek data, we estimated the H exponents corresponding to each one of the alleged cycles and tested their significance against the two random alternatives. The results are presented in columns 1 and 2 of the Table 4.11.

According to this evidence, only the smaller cycle has a significant Hurst exponent, as the low significance level indicates. The existence of this cycle is further confirmed by the flattening out of the V statistic curve after the 330-day period, as measured by its Hurst exponent in column 3 (points B to C) of Table 4.11. Hence, these first findings seem to support fractality and long-term dependence of the LSE series, although, the H estimate is quite low indicating very strong noise presence.

6 Alternative in the sense that both these alleged cycles cannot be true cycles since no flat (random walk) period is observed between them, as in the case of the Greek series.

R/S Analysis /LSE daily returns

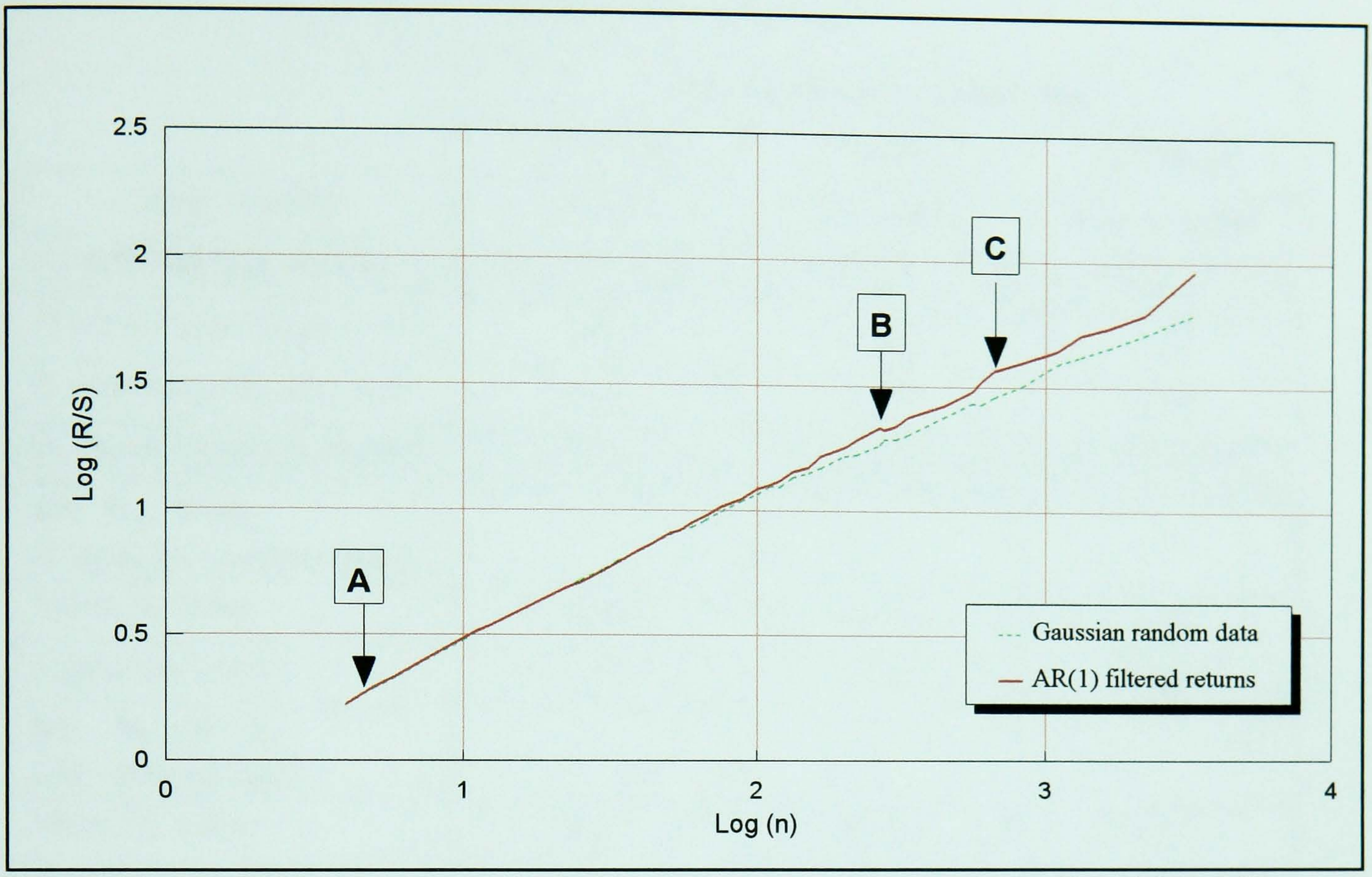

Figure 4.7a. Log(R/S) vs. Log(n) plot for the H exponent estimation of the AR(1) filtered daily LSE returns and a Gaussian random surrogate

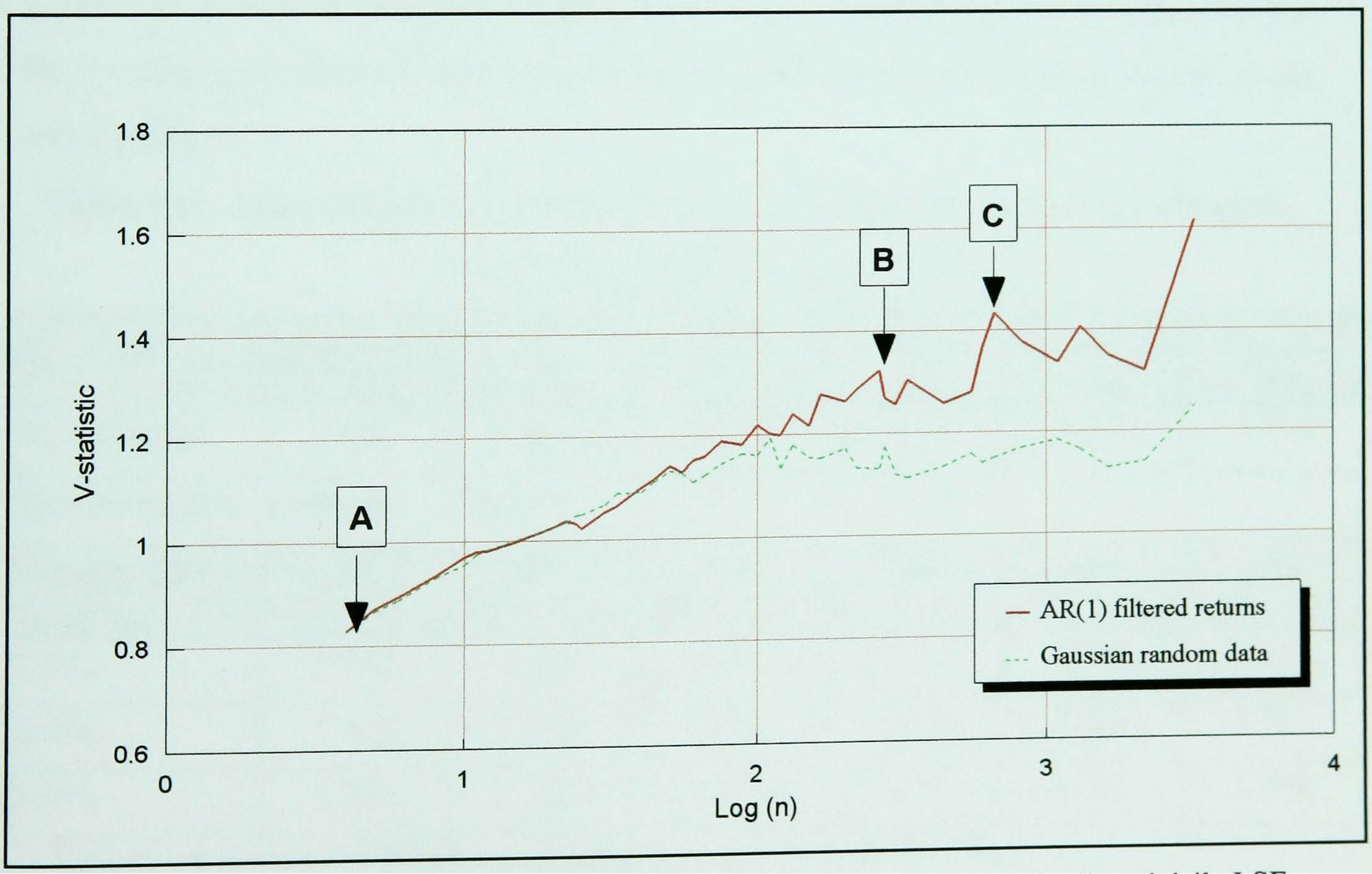

Figure 4.7b. V-statistic vs. Log(n) plot for the cycle length estimation of the AR(1) filtered daily LSE returns and a Gaussian random surrogate

Table 4.11 Hurst estimates and test of significance against two random alternatives of the daily LSE returns

Daily returns	Time period (in trading days) 1 (A to B)	2 (A to C)	3 (B to D)
	$1<n<330$	$1<n<660$	$331<n<3300$
A. R/S analysis results			
H estimate(original series)	0.63	0.62	0.57
H estimate(filtered series)	0.60	0.57	0.56
B. Bootstrapping results			
B1. H_0: $H=H_G$ (Gaussian random null)			
Mean H_G value	0.56	0.55	0.52
Significance level	0.012	0.121	0.283
B2. H_0 : $H=H_R$ (iid random null)			
Mean H_R value	0.57	0.55	0.52
Significance level	0.007	0.116	0.251

The above conclusions are questioned when the modified R/S statistic is employed. As Table 4.12 shows, the short-term dependence null is rejected for low q values but not for q values exceeding $n^{1/3}$ for both asymptotic and bootstrapped critical values of the test statistic.

Table 4.12 The modified R/S statistic of the LSE daily return series & bootstrapped critical values of the test statistic.

q	Andrew's 12	$N^{1/4}$ 9	$N^{1/3}$ 19	$N^{1/2}$ 82	100	150
V_q – statistic	1.933**	1.997**	1.835***	1.697	1.676	1.704
Bootstrapped Critical Values						
Significance level / q	**12**	**9**	**19**	**82**	**100**	**150**
1.0%	1.945	1.993	1.941	1.910	1.908	1.866
2.5%	1.815	1.846	1.808	1.802	1.800	1.790
5.0%	1.780	1.788	1.755	1.726	1.715	1.708

* Significance at one-tail 1.0% level according to the asymptotic critical values

** Significance at one-tail 2.5% level according to the asymptotic critical values

*** Significance at one-tail 5.0% level according to the asymptotic critical values

Although the findings of the above test seem ambiguous, since no indication about the proper q value exists, we can interpret them as a possibility for misleading results in our R/S analysis, due to autocorrelation bias.

To investigate this possibility further, we used a stronger AR(9) linear specification to filter our original LSE returns and repeated the R/S analysis. The results are presented in column 1 of the Table 4.13. and show that the Hurst exponent corresponding to the 330-one day cycle, drops to H = 0.56 a non-significant estimate against both random alternatives. Hence, the existence of this cycle is not verified and it seems to be an artifact due to linear autocorrelation.

Table 4.13 Hurst estimates and test of significance against two random alternatives of the AR(9) filtered one-day returns, the AR(1) filtered twenty-day returns and the AR(1) filtered one-day returns

	Time period (in trading days)		
	1 (daily data)	**2 (20-day data)**	**3 (daily data)**
	$1<n<330$	**$1<n<17$**	**$1<n<3300$**
A. R/S analysis results			
H estimate	0.56	0.60	0.56
B. Bootstrapping results			
B1. H_0: $H=H_G$ (Gaussian random null)			
Mean H_G value	0.56	0.64	0.53
Significance level	0.485	0.914	0.115
B2. H_0 : $H=H_R$ (iid random null)			
Mean H_R value	0.56	0.62	0.53
Significance level	0.371	0.885	0.108

However, taking into account that the modified R/S test has no power against autocorrelated chaotic processes as well as the possibility that a very strong linear filter might "wipe-out" part of the existing nonlinear or chaotic structure in a series [Brock (1986)], we further tested the hypothesis that the 330-day cycle is an indication of a noisy chaos structure.

If this is true, then the cycle should be independent of the time increment and a significant H estimate should appear for a 17-twenty day cycle corresponding approximately to the 330-one day cycle. To test this hypothesis a 20-day return series has been constructed from the original daily series and an AR(1) filter used to account for the autoccorelation found.

The results are shown in column 2 of the Table 4.13. The H estimate for the 17-twenty-day period is clearly insignificant, implying that the noisy chaos alternative is not supported.

Even if the LSE series exhibit no cycle, fractality, long-term memory (with no average cycle) and persistence of the series cannot be ruled out, unless the Hurst estimate of the total series is also tested against the random nulls. This is the case of a pure Hurst process i.e. a colored noise series of the FBM type, which might exhibit no-cycle but has a significant H exponent. We performed this test, too, and found (column 3 in Table 4.13) that the H exponent of the total AR(1) filtered series is not significant. Hence this alternative should be rejected, as well.

The above results are in accordance with Mills (1993), who finds no long-range dependence in monthly U.K. stock returns data. On the other hand, our findings do not support the existence of a 30-month cycle reported by Peters (1991a) for the U.K. market. However notice that Peters' application, besides suffering from technical problems (which he recognizes in more recent studies), does not account for autocorrelation which has been found in our analysis to bias the U.K. results.

In the technical level R/S analysis was proven to be a very powerful tool in detecting fractality, persistence and long-term dependence when enhanced with bootstrapping methodology to assess its statistical significance and combined with the modified R/S statistic to safeguard against serious autocorrelation bias.

Regarding our overall findings from the application of the R/S test, long-term dependence and fractality in the form of noisy chaos is supported for the ASE returns, while no indication of fractality or long term dependence, was found for the LSE series.

Chapter 5

APPLICATION OF THE CHAOTIC TECHNIQUES

5.1 VISUAL INSPECTION TECHNIQUES

A preliminary but often very useful step in chaotic analysis is a visual inspection of the data. Time series plots, phase-space plots and return maps provide different graphical representations of the series, which in pure chaotic systems can reveal the features of their deterministic structure.

In the case of noisy series such techniques do not offer much to the insight of a system's hidden structure.

Figure 5.1a-d shows the **time series plot,** i.e. the plot of the series observations X(t) as a function of time t, of the ASE and the LSE returns, a Gaussian random data and Lorenz attractor [Lorenz (1963)] one of the most famous chaotic systems generated by three differential equations[1]. Notice that the Gaussian random surrogate and Lorenz attractor are presented for comparison purposes, representing the extreme forms of a totally random (Gaussian white noise) and a purely chaotic system.

As we can see, in general and from a qualitative point of view, the two return series have a similar time series plot, which is quite distinct from that of Gaussian white noise and Lorenz series. However, the ASE series exhibit a less uniform plot than the LSE series, exhibiting more pronounced return movements in certain periods, indicating a possibly different overall structure.

Figure 5.1 (e-h), shows the two-dimensional **phase-space plot** of the above series which is constructed by plotting X(t) as a function of X(t-τ) where τ is the delay time parameter. Notice that the same plot can be constructed by plotting the derivative of X(t) against X(t). This method is directly related to the phase space reconstruction of a system described previously, according to which a chaotic system can be reconstructed

---

1 The Lorenz attractor is generated by the following system of first order differential equations: $dX/dt = \sigma\ (Y-X)$, $dY/dt = \rho X-Y-XZ$, $dZ/dt = XY-\beta Z$. A numerical solution can be achieved for parameter values $\sigma = 10$, $\rho = 28$, $\beta = 8/3$. This system which has been developed to describe a climatic attractor has a correlation dimension of 2.05.

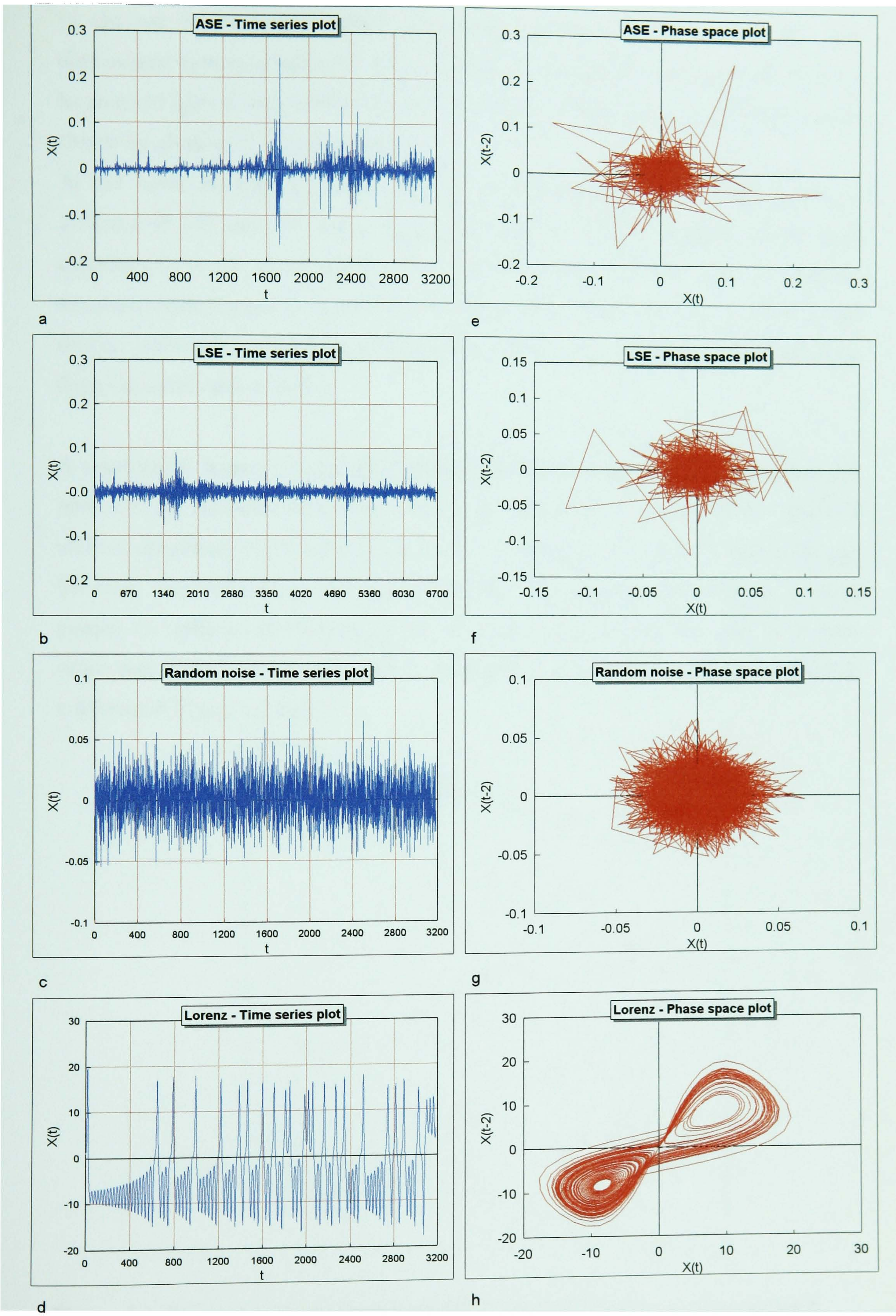

Figure 5.1a-h Time series plots and Phase Space plots of the ASE daily returns, the LSE daily returns, a Gaussian random noise data and Lorenz attractor

by the use of the lagged values of a single variable. So, in the case of a low dimensional system or when the data is periodic, a discernible pattern occurs. This can be seen in Figure 5.1(h), where a closed loop characterising the Lorenz system appears clearly by plotting X(t) against X(t-2).

At the other extreme, Gaussian random noise series [Figure 5.1(g))] shows no structure at all, and its phase space plot consists of a symmetric dense mass concentrated in the middle of the plot. The ASE and LSE series show no specific structure either (Figure 5.1 e and f respectively). Their phase space plots (for τ =2) are similar, indicating a strong noisy component in both series, but also, a different shape from Gaussian random series.

A third visual inspection technique is **Return Maps**. Return maps are theoretically related to the concepts of "surface of sections" and "Poincare map", which refer to a method introduced by Poincare at the end of last century. This method bridges the gap between continuous-time dynamical systems (flows) and discrete-time systems (maps), by replacing the flow of an nth-order continuous-time system with an (n-1)th-order map called the Poincare map. A graphical representation of the method is exhibited in Figure 5.2 below.

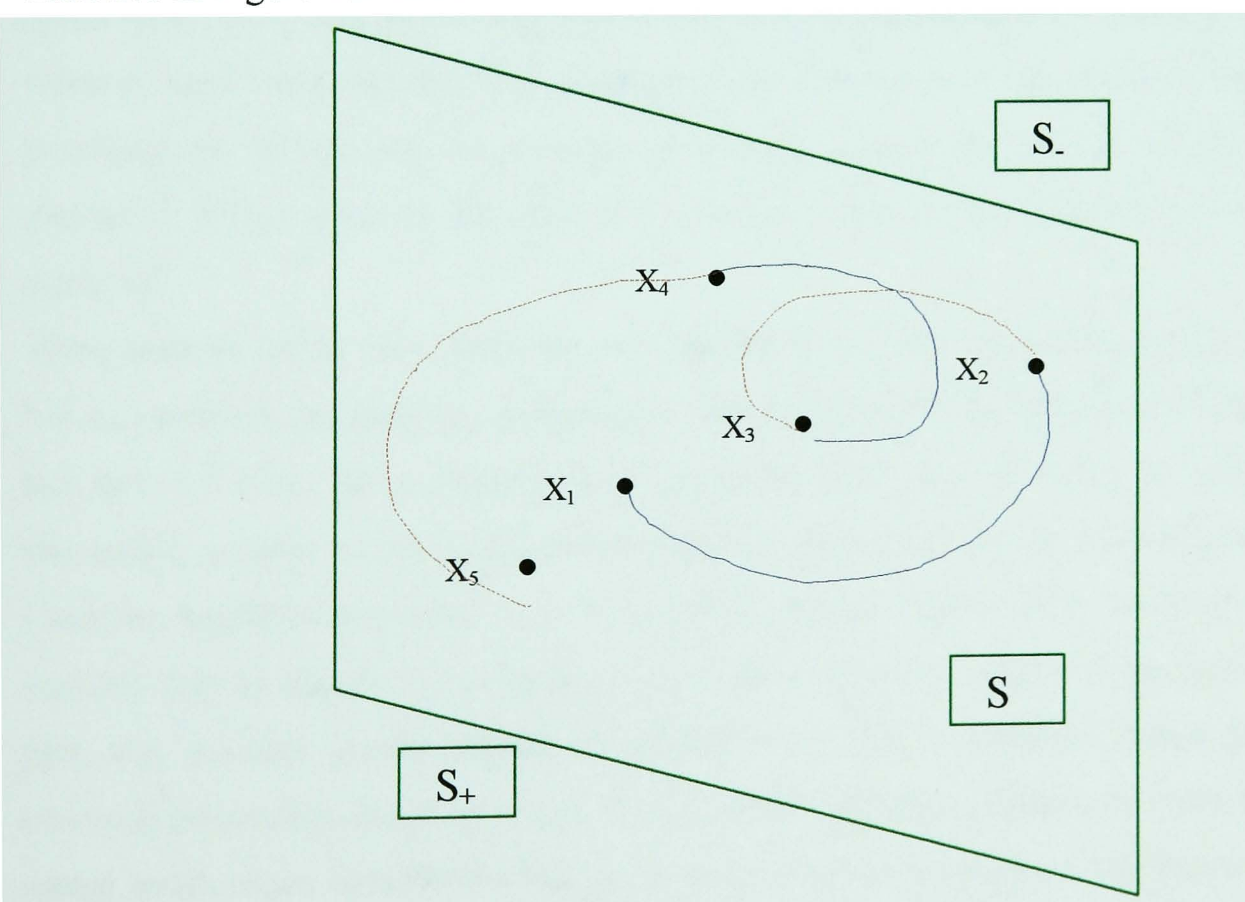

Figure 5.2 Poincare map of a typical trajectory $\varphi_t(x)$ intersecting a cross-section S.

Theoretically, the state space R^n is divided into two regions S_+ and S_- by an (n-1) dimensional hyperplane S transversal to a trajectory $\varphi_t(x)$. The latter will repeatedly pass through S, crossing from S_+ to S_- to S_+ etc. as in Figure 5.2. Given S we can define three different Poincare maps:

$P_+ : S \rightarrow S$. $P_+(x)$ is the point where $\varphi_t(x)$ first intersects S in a positive direction for t>0, and the sequence $\{x_1, x_3, x_5, \ldots\}$in fig. 5.2 is an orbit of the one-sided Poincare map P_+ .

$P_- : S \rightarrow S$. $P_-(x)$ is the point where $\varphi_t(x)$ first intersects S in a negative direction for t>0, and the sequence $\{x_2, x_4, \ldots\}$in fig. 5.2 is an orbit of the one-sided Poincare map P_- .

$P_{\pm} : S \rightarrow S$. $P_{\pm}(x)$ is the first point where $\varphi_t(x)$ intersects S in either direction for t>0, and the sequence $\{x_1, x_2, x_3, \ldots\}$in fig. 5.2 is an orbit of the two-sided Poincare map $P_{\pm}$, also called a first-return map.

Figures 5.3 a-d and 5.4 a-d, show the **Return Maps** of the four series presented above. The Return Map as a cross section of the phase plane, which reduces its dimension by one, can reveal the structure of a chaotic attractor better than a two-dimensional phase-space plot. For a real series, the return map can be constructed by plotting the data value at each local maximum or minimum versus its value at the previous maximum or minimum. In practice this is done by plotting X(t) at the time at which its first derivative X′(t) = 0 versus the value of X(t) at the previous time at which an extremum occurred.

There are two return map views for each series. In the first view, (Figures 5.3 a,c and 5.4 a,c) symbols are used to construct the plot, while in the second view, (Figures 5.3 b,d and 5.4 b,d), these symbols are connected with lines in order to reveal any discernible pattern. In our plots, a clear pattern occurs again for the Lorenz series. The Gaussian random data exhibits a symmetric pattern with a clear tendency of the variable X(t) to change sign under the return map. Both our return series show in this plot, too, a rather erratic pattern dominated by a noisy component which probably obscures an existing structure, if any. In all, visual inspection techniques show that our return series might be different than pure random noise, however, no further evidence is provided regarding their structure. They give the picture of a noise mixture data, the structure of which (if it exists) should be either high dimensional or obscured by noise.

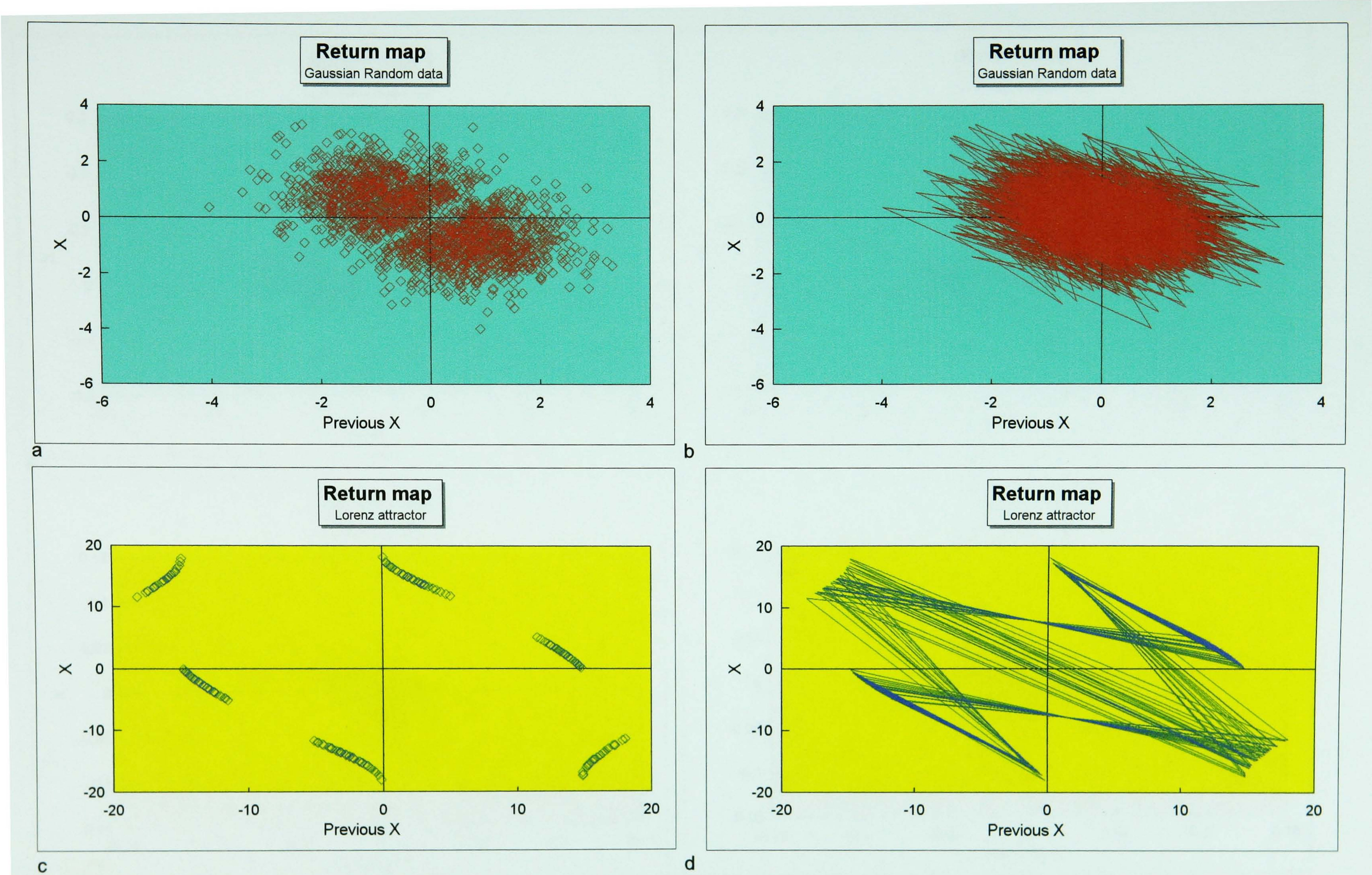

Figure 5.3a-d: a-b : Return map of a Gaussian random series. c-d : Return map of Lorenz attractor (Plots a and c are symbol plots while b and d are the corresponding line plots)

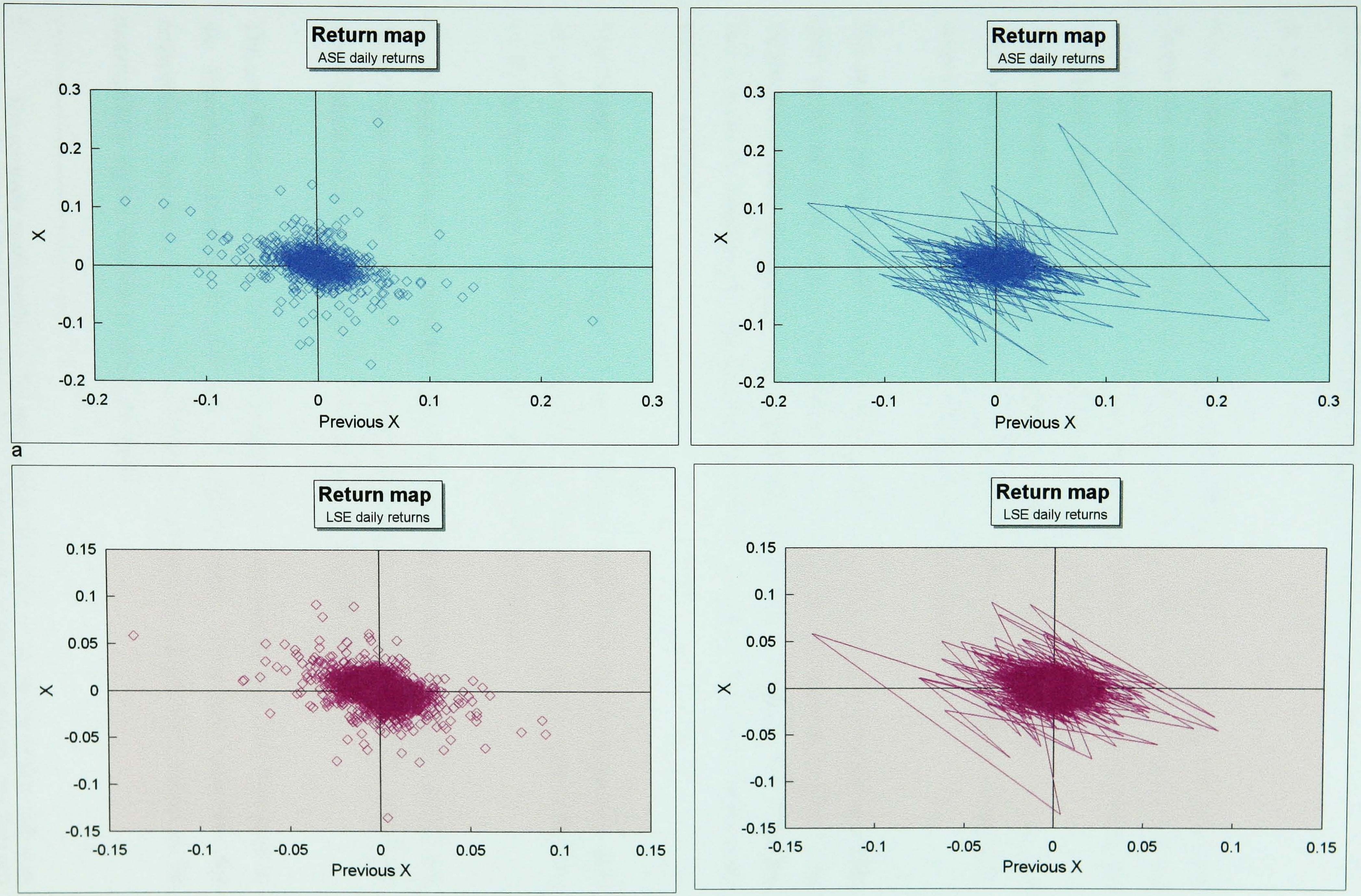

Figure 5.4a-d: a-b : Return map of ASE daily returns. c-d : Return map of LSE daily returns (Plots a and c are symbol plots while b and d are the corresponding line plots)

5.2 The Correlation Dimension Estimation

5.2.1 The ASE returns

The correlation dimension has been estimated by firstly following the approach frequently met in the Economics empirical literature. That is, we have calculated the correlation dimension for the raw data and then we tested the results by applying the "residuals" method and the "wing" and the "shuffle" diagnostics.
The correlation dimension of the ASE raw series was estimated for $m = 2,..,10$ and $\tau = 1$ and a saturating dimension $d \cong 6$ was found. This is shown in Figure 5.5-curve 1, where d is plotted against the range of embedding dimension m.

We applied the **"residuals"** method using an AR(2) specification to filter the raw ASE data as we did in the case of the BDS test. As we can see in Figure 5.5-curve 2, the dimension remains practically unchanged and saturation occurs again. This indicates that it is not the short-term temporal dependence that makes the dimension of the ASE series to saturate.

The **"wing"** diagnostic (Figure 5.5-curve 3) shows clearly that the Gaussian surrogate of our data performs as theory predicts, i.e. the dimension grows almost like m, verifying the (already established) non-normality of our series.

In the application of the **"shuffle"** diagnostic we followed the bootstrap method and 50 surrogate samples[2] of shuffled (randomised) residuals were constructed, having the same length and distribution as the AR residuals of the ASE series.

The correlation dimension has been computed for each sample and was compared to the dimension estimates for the original residual series. Dimension estimates for embedding dimensions $m > 6$ for all the 50 shuffled surrogates were found to be substantially higher than the original estimates.

---

2 The relatively small number of surrogate samples (50) used in the context of this method as compared to the 5000 randomised samples used in the R/S analysis. is due to the manual. pertinent and time consuming procedure in the calculation of the correlation dimension, for better accuracy reasons.

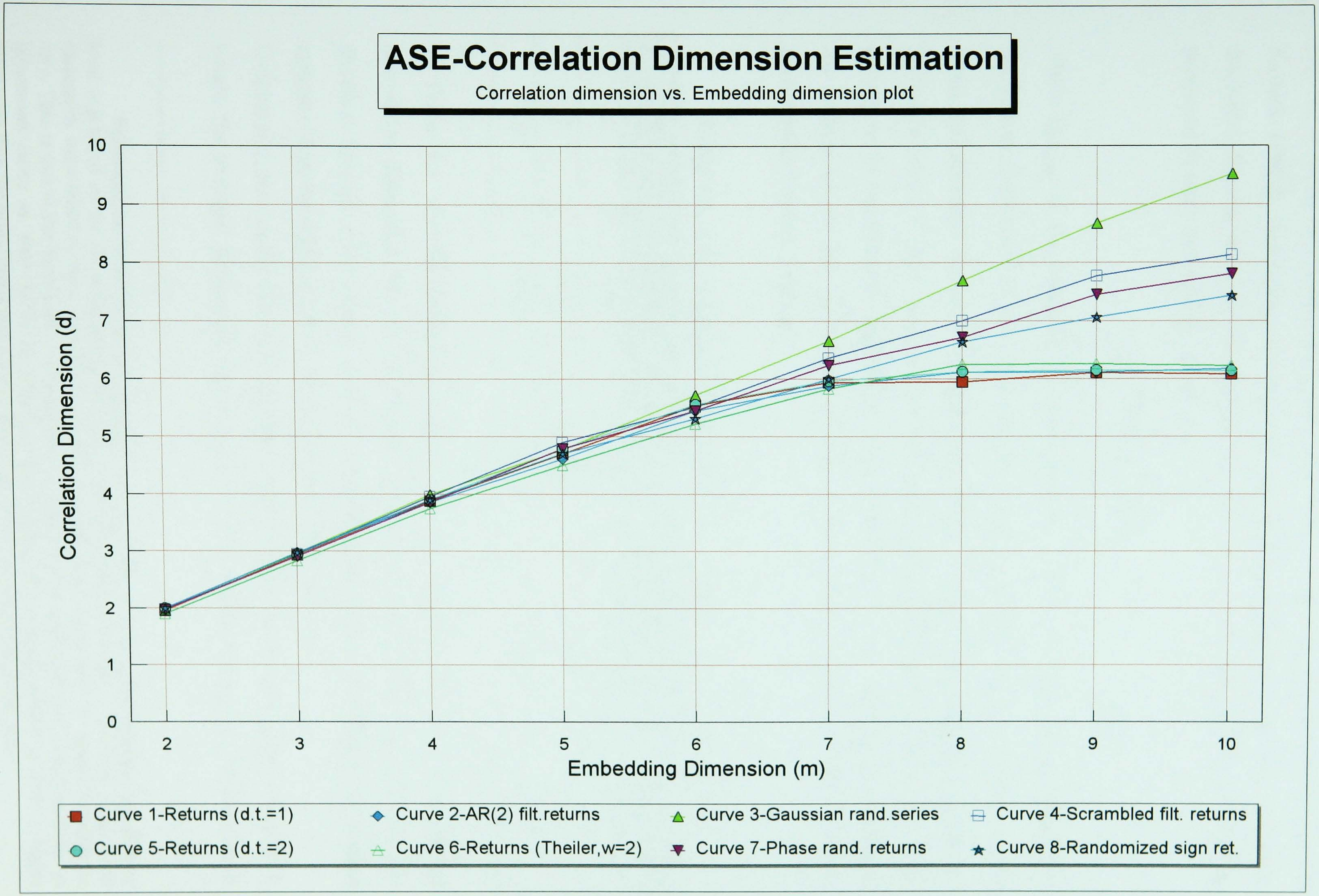

Figure 5.5 Plot of the correlation dimension estimate (d) versus different embedding dimansions (m = 1,..,10). The different curves correspond to various applications of the G-P algorithm to the original return series (curve 1,5,6) and to correlation estimates from different surrogate series (curves 2,3,4,7,8)

This means that the random null for the AR residuals should be rejected with a high empirical confidence level[3]. Due to the relatively low number of the shuffled surrogates used, we further tested this result using a method suggested by Theiler and Eubank (1993). Assuming normality of the sample distributions of the dimension estimates for the shuffled surrogates for each m>7, statistical significance can be measured in units of "sigmas" given by

$$S = \frac{\overline{Q}_{surrogate} - Q_{original}}{s_{surrogate}} \qquad (5.1)$$

where $Q_{original}$ is the value of the discriminating statistic for the original series (the correlation dimension of the ASE returns in our case) and $\overline{Q}_{surrogate}$ and $s_{surrogate}$ the mean and standard deviation respectively of the corresponding statistics for the 50 surrogate series. S values greater than 1.96 "sigmas" indicate that the dimension of the surrogates is significantly greater (at a 5% significance level) than the dimension of the original series. The statistical assessment of the shuffle test for the ASE series is presented in Table 5.1 below.

Table 5.1 ASE: Statistical assessment of the "shuffle" diagnostic results

m	$Q_{original}$	$\overline{Q}_{surrogate}$	$s_{surrogate}$	S
8	6.11	7.01	0.1075	8.37
9	6.12	7.77	0.1170	14.10
10	6.16	8.13	0.1162	16.95

In Table 5.1, the discriminating statistic for the original series is the estimated correlation dimension for the AR filtered series. It is obvious from Table 5.1 that the shuffled filtered ASE returns exhibit a significantly higher correlation dimension estimate than the original series, as the high S values indicate.

Graphically, the results from the "shuffle" diagnostic are shown in Figure 5.5-curve 4, where the average dimension value from the 50 shuffled surrogates for each m is

---

3 Following the methodology presented in R/S analysis for estimating the empirical significance level (s.l) based on the 50 surrogates constructed by randomisation, we found s.l = 0.019, although all surrogates were found to have a higher dimension than the actual series and s.l. should normally be zero. This is due to the ratio used to estimate s.l., as suggested by Noreen (1989), which includes an adjustment factor as: (ns+1)/(NS+1), where ns=the number of surrogates found to have a higher dimension than the actual series and NS=the total number of surrogates used in the test.

plotted against the dimension of the embedding space. We can clearly see that the shuffled series' dimension curve is not saturating and is distinctively above the saturating curve of the original series.

Our next step was to use a number of approaches currently followed in the Natural Sciences literature, in order to avoid possible shortcomings of the preceding analysis. Specifically, Mayfield and Mizrach (1991) argue that the "residuals" method has little power against dependence in higher moments and propose the application of the method of delays by the use of the proper delay time[4].

We repeated our dimension calculations with delay time $\tau = 2$ which corresponds to the first zero crossing of the autocorrelation function of our series. Our estimates were found to be robust to the delay time reconstruction (Figure 5.5-curve 5) and a saturating dimension $d \cong 6$ occurred again.

Theiler (1986) suggests a modification of the correlation integral which reduces the temporal correlations between nearby points and bases the dimension estimate on the spatial correlations which are the ones to reflect the geometrical features of the (supposed) attractor. We took this precaution and reproduced our calculations for a cut-off parameter equal to the decorrelation time (W=2). Our dimension estimates once more did not change (Figure 5.5-curve 6).

In addition, we have applied the **"phase randomisation"** technique [Theiler (1991), Provenzale et. al. (1992)] having power against linear correlation, as well as, against different kinds of autocorrelated processes like fractal noises with a $1/f^{a}$ spectrum and non-linear multifractal stochastic processes. Since R/S analysis favoured a noisy chaos alternative against fractal noise, this technique could also help to verify this finding.
Recall that "phase randomisation" entails the construction of surrogate data, which is random but has the same variance and autocorrelation as the raw data. This is done by taking the Fourier transform of the original data, randomise the phases and take the inverse Fourier transform. If the dimension results are found to be different from the

---

4 Almost all applications in the Economics' literature use a linear filter to remove temporal correlation and a delay time $\tau = 1$.

original series, the null of a correlated random process can be rejected in the context of a bootstrap application.

Following once again the methodology described in Chapter 2, randomised phase surrogates were constructed. The correlation dimension was estimated for each surrogate (for m=2,..,10). In all cases, the correlation dimension was non-saturating and for m>6, it was found increasingly higher than the saturating limit $d \cong 6$.

This finding was further assessed by the same methodology used in the case of the "shuffle" diagnostic. The results are presented in Table 5.2.

Table 5.2 ASE: Statistical assessment of the "randomised phase" results

m	$Q_{original}$	$\overline{Q}_{surrogate}$	$s_{surrogate}$	S
8	6.13	6.72	0.0583	10.12
9	6.15	7.45	0.0354	36.72
10	6.14	7.81	0.0404	41.34

This time $Q_{original}$ represents the correlation dimension estimates for the ASE returns with delay time $\tau = 2$.

The results of this test are clearly rejecting the autocorrelated noise null at a very high significance level, as the S values indicate, and the findings of the R/S analysis, which have also rejected a fractal noise alternative in favour of noisy chaos, are supported.

This is pictured in curve 7 of Figure 5.5 where the average d of the 50 randomised phase surrogates for each different m is plotted against m. As in the case of the shuffled series' surrogates, the randomised phase dimension curve is not saturating and lies above the original series' dimension curve.

Finally, we investigated whether the saturating dimension results in our ASE series are solely due to its variance signature. This has not be done before in the literature, however, it is especially important when we deal with stock market returns for which several specifications modelling the dependence in variance (ARCH-type models) have been suggested.

We call this approach the "**randomised sign**" technique, since in order to construct a surrogate having the same variance but a different mean than the original series, we randomise its signs. The null tested via the surrogate data method and the bootstrap is that the correlation dimension estimate for the original series is due solely to its variance signature. Once the dimension estimate of the surrogates is found significantly different than the dimension for the original series, the null can be rejected.

As in the previous tests, 50 randomised sign surrogates have been constructed. The correlation dimension of all these surrogates was found to be non-saturating and increasingly higher than 6 for m>8. This is shown in curve 8 of Figure 5.5, where the average d of the 50 surrogates for each m is plotted. Curve 8 lies a little lower than the corresponding ones to the scrambled and the randomised phase curves. However, it is clearly not saturating and higher than the saturating limit, picturing the rejection of the null hypothesis.

This is also shown in Table 5.3 where the statistical significance of the difference between the average dimension for the randomised sign surrogates and the dimension of the original series is assessed. The high S values indicate in this case too, significantly higher dimension estimates for the surrogate series, rejecting the null.

Table 5.3 ASE: Statistical assessment of the "randomised sign" results

m	$Q_{original}$	$\overline{Q}_{surrogate}$	$S_{surrogate}$	S
8	6.13	6.64	0.0552	9.24
9	6.15	7.06	0.0720	12.64
10	6.14	7.43	0.0683	18.89

In conclusion, the ASE series saturate at a rather high correlation dimension (after taking into account different kinds of technical issues) and shows clearly different qualitative behaviour from Gaussian random or other kinds of random processes. Moreover, our dimension estimate is similar to the ones reported by Sheinkman and LeBaron (1989) for US stock market data and by Vaidyanathan and Krehbiel (1993) for US futures series, a fact which could be interpreted as an indication of structural similarities between markets with different characteristics.

In technical terms, these results suggest that a deterministic explanation for the ASE series cannot be ruled out, however, a simple low-dimensional attractor is highly unlikely unless it is obscured by a large amount of noise. We should also be very cautious in claiming that even a high-dimensional attractor exists. Our saturating dimension value is too high to be fully supported by the length of our series, according to the less conservative approaches for data requirements discussed in Chapter 2. Saturation "per se" might be an artefact due to small sample bias [Ruelle, (1990), Ramsey et. al. (1990)]. Lack of strict stationarity, very common to economic series, might also undermine our findings, although this issue has been addressed using as a tool the BDS test.

Nevertheless, from a qualitative point of view, even if the estimated correlation dimension is not the correct one, the limited downward bias of the Gaussian random version of our data (Figure 5.5-curve 3) could be considered as an optimistic indication for reasonable bias at this data length. In addition we believe that the non-contradictory results of the different chaotic techniques employed, enhanced by the bootstrap methodology, can reliably support the view that the qualitative behaviour of our series is quite distinct from randomness. In this case, the opposite problem, i.e. the possibility of an upward bias in the dimension due to the small size of our data, cannot be ruled out.

The remaining question to be answered is related to noise and the effect it might have on our dimension estimates. This issue is addressed in the next Chapter.

5.2.2 The LSE returns

The correlation dimension estimation of the LSE raw series for m = 2,..,10 and $\tau = 1$ is shown in curve 1 of Figure 5.6. This time, unlike the ASE series case, no saturation occurs and the correlation dimension is close to 9 for m=10, indicating that the behaviour of the LSE series is indistinguishable from that of a random series.
The **"residuals"** method was also applied to see if the filtering of the linear autocorrelation present in this data by an AR(1) filter would have any effect on our dimension estimates. As we can see in Figure 5.6-curve 2, the dimension remains practically unchanged and a non-saturating behaviour of the dimension curve is apparent once more.
In the application of the **"shuffle"** diagnostic we followed exactly the same procedure described previously in the case of the ASE series and 50 surrogate samples of shuffled residuals were constructed. The correlation dimension of these surrogates for each m was found to be practically identical to the dimension estimate of the raw and the filtered LSE series. Actually, unlike the case of the ASE series where the empirical significance level of the bootstrap methodology was close to zero, almost half of the LSE shuffled series for each embedding dimension were found to have lower dimension than the original series[5]. The statistical assessment of the "shuffle" diagnostic for the LSE series is shown in Table 5.4

Table 5.4 LSE: Statistical assessment of the "shuffle" test results

m	$Q_{original}$	$\overline{Q}_{surrogate}$	$s_{surrogate}$	S
8	7.30	7.42	0.0824	1.456
9	7.98	8.13	0.0892	1.682
10	8.47	8.59	0.0935	1.283

The small S values (<2), indicate non-significant difference between the discriminating statistic (correlation dimension) for the surrogates and the original series. According to these results, the null of the test is not rejected and the LSE series

[5] Empirical significance levels range from 0.46 to 0.58 for m>5.

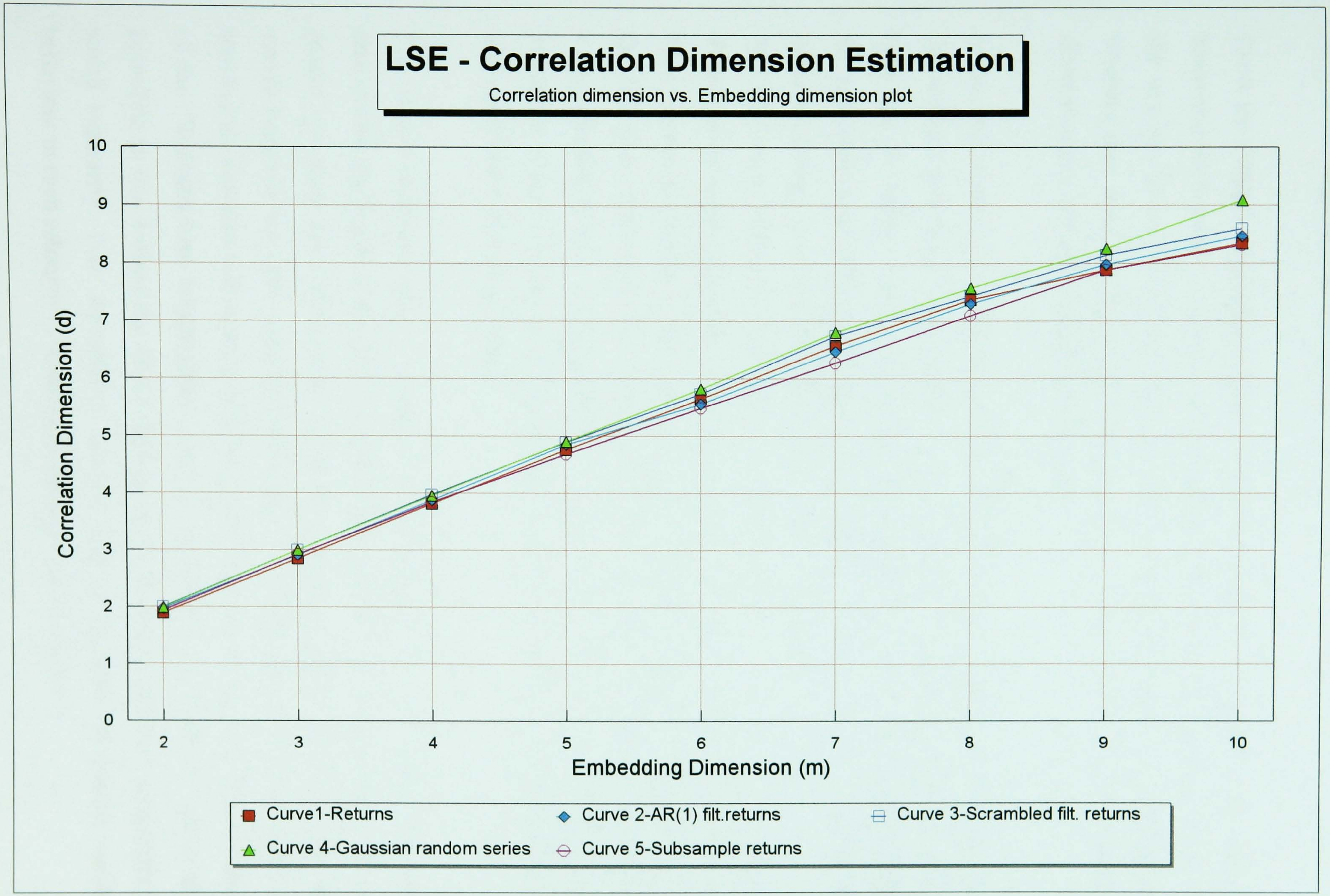

Figure 5.6 Plot of the correlation dimension estimate (d) versus different embedding dimensions (m = 1,..,10). The different curves correspond to the application of the G-P algorithm to the original return series (curve 1) and to correlation estimates from different surrogate series (curves 2,3,4,5)

can not be distinguished from random iid series having the same distributional characteristics. Graphically, this is shown in Figure 5.6 where the average dimension estimate of the 50 surrogates for each m is projected in curve 3. As we can see, the latter is indistinguishable from curves 1 and 2.

Even the "**wing**" diagnostic (Figure 5.6-curve 4) shows no qualitatively different behaviour between the Gaussian random and the LSE series. In fact, only for m=10, the average dimension of the Gaussian random surrogates gives a clearly higher estimate than the LSE series, reflecting probably, and in terms of the correlation dimension test, the non-normality of the latter.

In the case of the LSE series, there is no need to proceed further with the rest of the "surrogate" data based tests related to the correlation dimension estimation. These tests can be useful when a saturating dimension has been found and different surrogates are employed to investigate whether this is an artifact due to several causes (autocorrelation, variance signature etc.) that bias the dimension downward.
The LSE series show no indication of such a saturation point and its correlation dimension is indistinguishable from the one exhibited by a random series. However, there are two remaining issues still unresolved. The first is the noise problem. As in the case of the ASE series, there is a possibility of an underlying chaoticity masked by a large amount of noise. Although the R/S analysis, which is robust to noise, shows no indication of such a structure in the case of the LSE series, the problem remains and is addressed in the following Chapter.

The second important issue is related to the data length. The LSE series have more than double the length of the ASE series and show no saturation which means that our results are much more robust than in the previous case. However, we can use the excess length of the LSE series to test whether we would have a different result, in terms of the correlation dimension test, if a smaller data should be used. This is a form of the **"Independent Realisations"** method presented in Chapter 2. Recall that according to this method the data is divided into subsamples and each subsample is tested separately. Low dimensional chaotic data is expected to exhibit similar behaviour in each subsample, verifying in parallel its stationarity.

Due to the small number of observations, we could not apply this test in the case of the ASE data. Yet, this is possible with the LSE series. To this end, we have considered a sub-sample of the latter, spanning exactly the same period as the ASE series, having a length of 3364 observations similar to the ASE length of 3181 observations. This new series is also AR(1) autocorrelated and its correlation dimension was estimated with $\tau=1$, for m=2,..,10.

The results are shown in Figure 5.6-curve 5 and no difference is observed in the qualitative behavior of the correlation dimension curve. No saturation is established and the dimension for each m is almost identical to the corresponding to the longer series. Hence, we can claim that our conclusions concerning the LSE data are not sensitive to the data length.

This test supports also – though in an indirect way – the reliability of our results concerning the ASE series, since we do not have any reason to believe that in the case of the ASE series it is the data length that affects the dimension estimates, when we have proved that this is not happening with the LSE series.

5.3 THE LARGEST LYAPUNOV EXPONENT ESTIMATION

The preceding analysis does not rule out a chaotic explanation for the ASE series; so, in this section, we first see whether the Largest Lyapunov Exponent (LLE) estimation supports the above findings as well. The LLE (L_1) estimation for the ASE series requires the definition of some basic parameters, as described in Chapter 2.

Since the dimensionality of the ASE returns was found to be approximately d = 6, the embedding dimension was fixed to m = 13, according to the $m \geq 2d+1$ relationship, defining the proper embedding [Takens (1980)]. For this embedding dimension, the delay time should be fixed to $\tau = 12$, following the $Q = m\tau$ rule [Wolf et. al. (1985), Peters (1991a)], where the orbital period Q = 155 days corresponds to the cycle-length found by R/S analysis in the previous Chapter. The maximum length of growth between the two points was set to SCALMAX = 0.1 of the data range and the accuracy level SCALMIN = 0.01 of the data range.

To extract some useful information from the behaviour of LLE, we considered its dependence on certain parameters such as the evolution time (t_s) and the embedding dimension (m). By varying t_s, we can test the stability of L_1. By varying m we can test whether its behaviour is consistent to that of a chaotic system [Bountis et. al., (1993)]. According to Abarbanel et. al. (1990), a chaotic system is expected to exhibit a higher (lower) positive LLE when the dimensionality of the embedding space is reduced (increased) since the attractor occupies a larger (smaller) portion of the available space On the other hand, the LLE behaviour for a random system is quite different since L_1 remains practically unchanged for different embedding dimensions as a random series tends to spread out uniformly in the available space.

The estimation of L_1 from the ASE series for the fixed parameters noted above and varying evolution time t_s = 5, 10,15 and 20 is shown in Figure 5.7. In all cases, L_1 is positive and exhibits a stable convergence as expected from a chaotic system. In addition, the variability of the Lyapunov estimate is decreasing with increasing evolution time and stabilises for t_s = 15 (The L_1 estimate for t_s = 20 is almost the same with the corresponding estimate for t_s = 15).

This property is in accordance with sensitivity to initial conditions of a system evolving on a strange attractor structure.

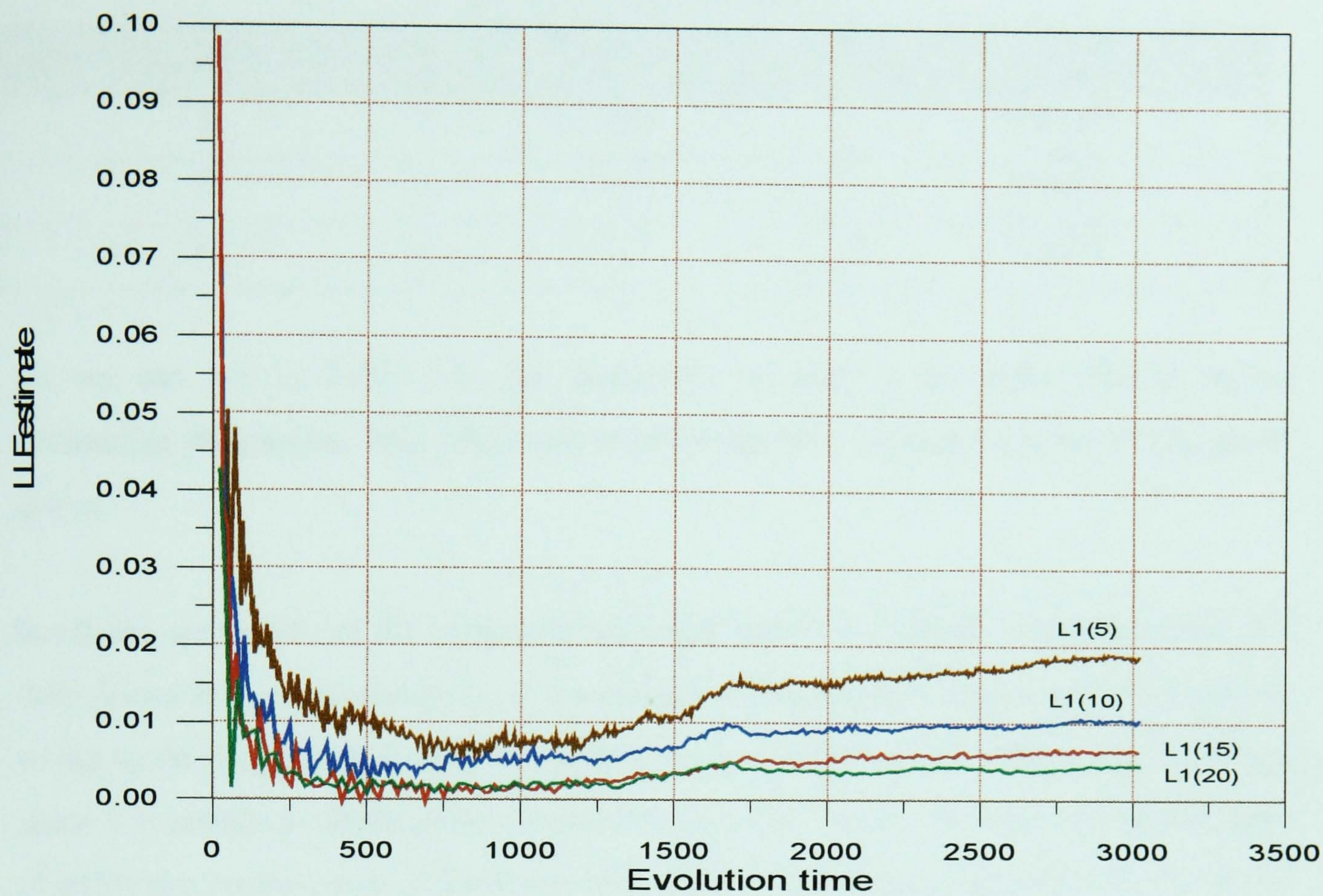

Figure 5.7. ASE returns: convergence of the Largest Lyapunov Exponent for varying evolution time $t_s = 5,10,15, 20$.

According to this principle, the displacement of two nearby orbits quickly increases for small values of the evolution time step, while, in contrast, for large values, the trajectories on the strange attractor converge and the estimated L_1 values are low [Pavlos et. al. (1994)].

For the stabilising value of $t_s = 15$ and the fixed parameters mentioned above, $L_1 = 0.0063$ bit/day. This means that if initial conditions could be measured with precision of one bit, memory (or predictive ability) would be lost after $1/0.0063 \cong 155$ days which coincides with the decorrelation time (or cycle) found by R/S analysis, implying dissipation of the long memory effect in the series.

To investigate the behaviour of the LLE for changing embedding dimension, we kept the evolution parameter constant to $t_s = 15$ and varied the m parameter (m = 6, 9, 13). According to the Q = mτ rule, τ will also vary as displayed in Table 5.5.

Table 5.5 LLE estimation for the ASE returns,

($t_s = 15$ and $m = 6, 9, 13$)

m	τ	L_1
6	26	0.0153
9	17	0.0089
13	12	0.0063

As we can see in Table 5.5, the Lyapunov estimate is inversely related to the embedding dimension. This behaviour is expected from a chaotic system as discussed above.

In all, the estimation of the Lyapunov exponent supports a chaotic explanation for the ASE series and seems also to be in line with the findings of the R/S test, with respect to the cycle length of the data. However, this test suffers from serious shortcomings since it is unable to distinguish between random and chaotic alternatives on the basis of either the positiveness or the behaviour of the LLE estimate for changing m, as we can see in Table 5.6.

Table 5.6 LLE for different chaotic, random and periodic specifications

		Henon			Lorenz			Gaussian random	Fractal random	Sine
m	t_s	L_1	m	t_s	L_1	M	t_s	L_1	L_1	L_1
2	2	0.603	3	3	0.071	3	3	0.609	0.325	-0.00015
3	2	0.559	5	3	0.034	6	3	0.147	0.059	-0.00005
5	2	0.558	7	3	0.011	9	3	0.093	0.035	-0.00002
2	5	0.513	3	5	0.055	3	5	0.350	0.207	-0.00004
3	5	0.423	5	5	0.030	6	5	0.121	0.045	-0.00003
5	5	0.413	7	5	0.006	9	5	0.085	0.034	-0.00002
2	7	0.399	3	7	0.036	3	7	0.264	0.199	-0.00004
3	7	0.362	5	7	0.013	6	7	0.093	0.044	-0.00003
5	7	0.388	7	7	0.003	9	7	0.070	0.030	-0.00002

Table 5.6 shows the LLE estimates (for different m and t_s parameters) for two chaotic, two random and one periodic specification, namely, Henon map, Lorenz attractor, a Gaussian random series, a fractal ($1/f^a$) random series and a sinusoid[6] data respectively. Notice that m values for the two chaotic specifications have been chosen to include a value close to their correlation dimension, a value to satisfy the $m \geq 2d+1$ rule and an intermediate value.

As we can see, the Lyapunov exponent test can distinguish between periodic and non-periodic data since in the first case the estimated LLE is always negative, irrespective of the parameters employed. However, both random and chaotic series exhibit positive LLE estimates, which are inversely related to the increase in the embedding dimension. Hence none of the criteria mentioned above can be used to distinguish between random and chaotic alternatives.

The same results hold when we compare the LLE estimates for the ASE and LSE series. Table 5.7 shows the LLE estimates for the ASE and LSE returns, the shuffled

Table 5.7 LLE for the ASE and LSE original and shuffled series

		ASE returns	ASE shuffled returns	LSE returns	LSE shuffled returns	LSE returns (small set)
m	t_s	L_1	L_1	L_1	L_1	L_1
6	5	0.0284	0.0805	0.0852	0.1294	0.1006
9	5	0.0213	0.0518	0.0458	0.0499	0.0250
13	5	0.0122	0.0158	0.0145	0.0168	0.0141
6	10	0.0194	0.0376	0.0412	0.0590	0.0453
9	10	0.0130	0.0231	0.0220	0.0193	0.0155
13	10	0.0061	0.0075	0.0059	0.0092	0.0057
6	15	0.0145	0.0350	0.0325	0.0482	0.0372
9	15	0.0099	0.0218	0.0189	0.0174	0.0118
13	15	0.0053	0.0046	0.0075	0.0074	0.0053

[6] This data is generated by a simple periodic function sin(t/10), at integer values of t>0.

ASE and LSE returns and the "small" LSE series, which span the same period as the ASE series. The m and t_s parameters are the same as the ones used above to calculate the LLE for the ASE series. This time, for comparison purposes, we use the same delay time ($\tau = 2$) for all the series tested, and SCALMAX and SCALMIN are defined as portions of the data range (0,1 and 0,01 respectively), as described in the previous section.

As Table 5.7 shows, all series exhibit positive LLE which vary with m and t_s parameters in a way expected from chaotic specifications. In addition, the magnitude of the exponents is similar between the original and the shuffled series for both data sets. These results, too, support the conclusion that neither the behaviour of the LLE or any other criterion can be used to distinguish between chaotic and random specifications through the Lyapunov exponent estimation, at least when the Wolf et. al. (1985) algorithm is used.

In all, the results from the Lyapunov exponent estimation do not rule out a chaotic explanation for the ASE returns. On the contrary, the results here corroborate with the results from the R/S test with respect to the cycle of the series. However, we have shown that unlike the correlation dimension estimation, the LLE approach can not be used on its own to distinguish between chaotic and random alternatives.

Chapter 6

NOISE FILTERING AND TESTING OF NOISE FILTERED SERIES

6.1 NOISE FILTERING BY THE SVD METHOD

6.1.1 Overview and implementation issues

In Chapter 1, the theoretical aspects of the SVD method as an alternative in phase space reconstruction of a dynamical system have been presented analytically. Here we shall focus on the ability of this method to filter away noise, thus providing a very useful tool in analyzing more efficiently noisy data like the ASE and LSE series.

Recall that the essence of the method is to generate a new projection basis (coordinate system) for the trajectory matrix **X**. This can be done efficiently by diagonalizing the covariance matrix **V** whose elements are the covariances of the observations forming the χ_i vectors:

$$\mathbf{V} = \mathbf{X}^{\mathrm{T}}\mathbf{X} = \frac{1}{N}\sum_{i=1}^{N}\chi_i\chi_i^{\mathrm{T}} \qquad (6.1)$$

The diagonalization and decomposition of **V** produces a set of eigenvectors (called the singular vectors of **X**) which form an (mxm) orthogonal matrix consisting of the projection basis of **X**, and a corresponding set of m-eigenvalues (whose square roots are called the singular values of **X**) which define the rank of **V**, that is, the number of the linearly independent vectors that can be constructed from the trajectory. The product of the projection is a new (Nxm) matrix the columns of which (called principal components or singular functions) correspond to the delayed column vectors of the trajectory matrix.

In the case of a noise free system, the number of the d positive eigenvalues of **V** correspond to the dimensionality of the embedding subspace explored by the trajectory, while the remaining ones (m-d) are equal to zero. Hence, dimensionality is reduced to an eigenvalue problem.

However, if noise is present, all the eigenvalues are non-zero and only the number of the dominant ones gives the "statistical dimension" [Vautard and Ghil, (1989), Medio

(1992)], that is, a reliable upper bound to the dimensionality of the subspace explored by the deterministic part of the trajectory.

Once the dominant eigenvalues have been identified, noise filtering is possible, since a "reduced" trajectory matrix can be constructed from the projection of **X** onto the reduced eigenvector basis that corresponds to the significant eigenvalues of **V**

The empirical problem is how to identify the important eigenvalues in each case. According to Broomhead and King (1986a,b) for the proper embedding parameters[1] noise should create a noise floor identifiable in the plot of the eigenspectrum and the dominant (emerging) eigenvalues are considered to be the ones above this noise floor.

Alternatively, Broomhead, Jones and King (1987) and Medio (1992) suggest that the emerging eigenvalues can be identified by considering the shape of the corresponding eigenvectors. Theoretically, the emerging ones will have a rather regular shape close to that of orthogonal polynomials of degree i - 1, where i is the order of the eigenvector.

An important issue is how to assign statistical confidence to the eigenvalues in eigenvector problems. Vautard et. al. (1992) and Elsner and Tsonis (1996) stress the difficulty of solving this problem. Two feasible suggestions can be found in the literature. Under the assumption that the eigenvalues are normally distributed, Vautard and Ghil (1989) provide a formula adapted by Ghil and Mo (1991) according to which the 95% confidence interval of an eigenvalue λ_κ is given by the variance formula $\lambda_\kappa \pm \lambda_\kappa(2/N_d)^{1/2}$ where N_d is the number of degrees of freedom for a given embedding dimension m, estimated as $N_d=(N_T/m)-1$, N_T being the sample length. Vautard et. al. (1992) indicate that this formula is quite conservative and suggest a parametric bootstrap alternative. In the context of the latter, 100 Gaussian random surrogates of the original sample are created and the 95% error bars are calculated as $\lambda_\kappa \pm 1.96\sigma_\kappa$, where σ_κ is the std for the k-th eigenvalue esitmated with the 100 realizations of λ_κ.

In our application we have used both approaches to estimate the 95% confidence intervals of our eigenvalue estimates. The differences found were trivial so we report (in the form of error bars to the eigenvalue spectrum) only the parametric bootstrap

---

1 The dimensionality of the (mxm) covariance matrix.

estimates produced from 1000 Gaussian random surrogates of our series having the same length, mean and variance.

6.1.2 Empirical Evidence

6.1.2.1 The case of the ASE series

We applied the SVD method to our ASE return series for different embedding dimensions (m = 5, 8, 11, 15). Figure 6.1 shows the eigenspectrum plot for the return series. For m=15 a noise floor is apparent for m>11, with 9 eigenvalues to emerge above the noise floor, but with 6 of them to look as the dominant ones.

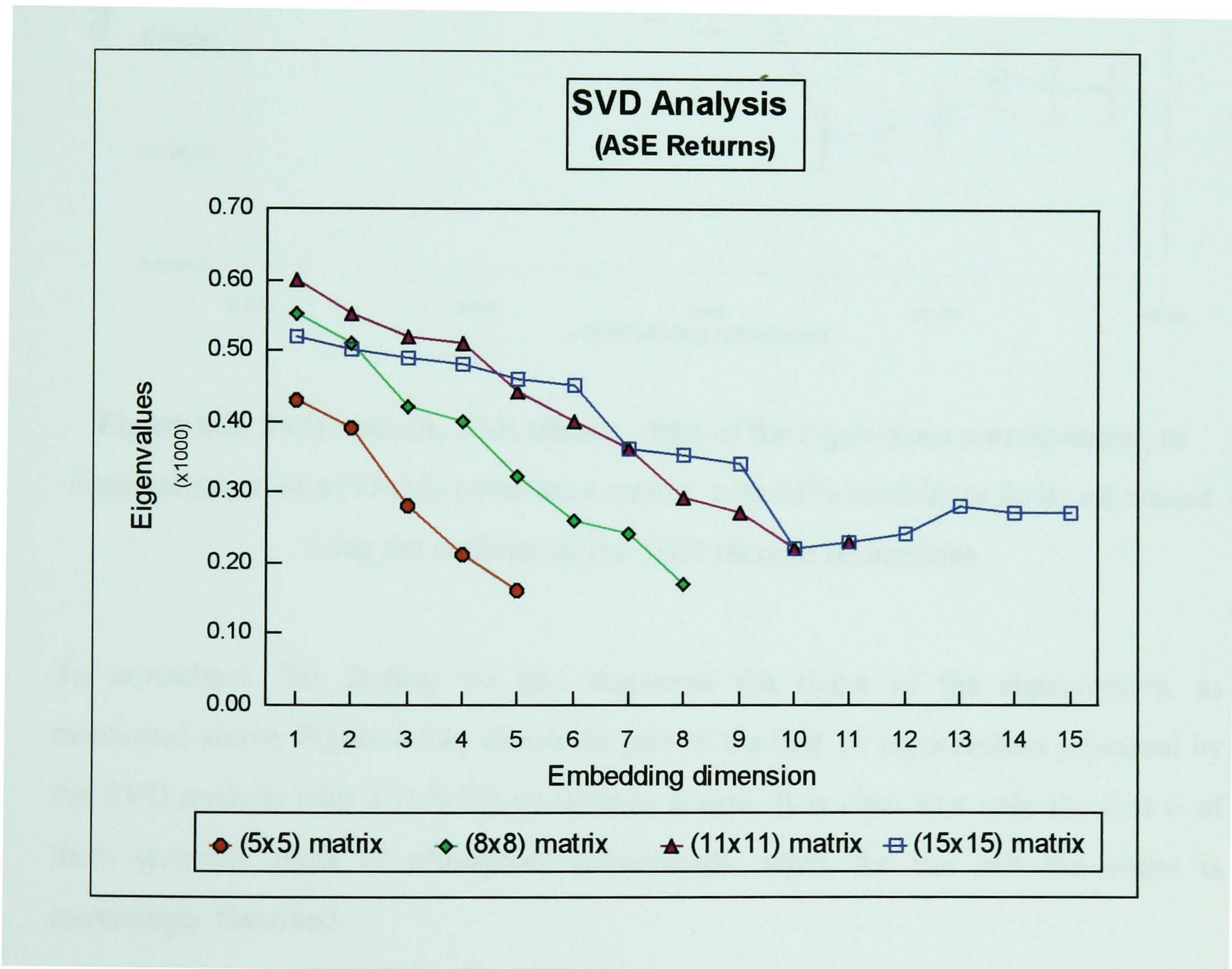

Figure 6.1. SVD analysis - Plot of the eigenvalues versus embedding dimension for the ASE returns. The four curves correspond to diagonalization of a (5x5), (8x8),(11x11) and (15x15) covariance matrix respectively

Figure 6.2 shows the eigenspectrum for m=15 with error bars indicating the 95% confidence interval estimated by bootstrapping as described above. Our prior conclusion that 6 dominant eigenvalues are observed, is verified. Taking into consideration the confidence intervals, the last 9 eigenvalues cannot be significantly distinguished from the noise floor. Hence only the first six can be considered as the significant ones.

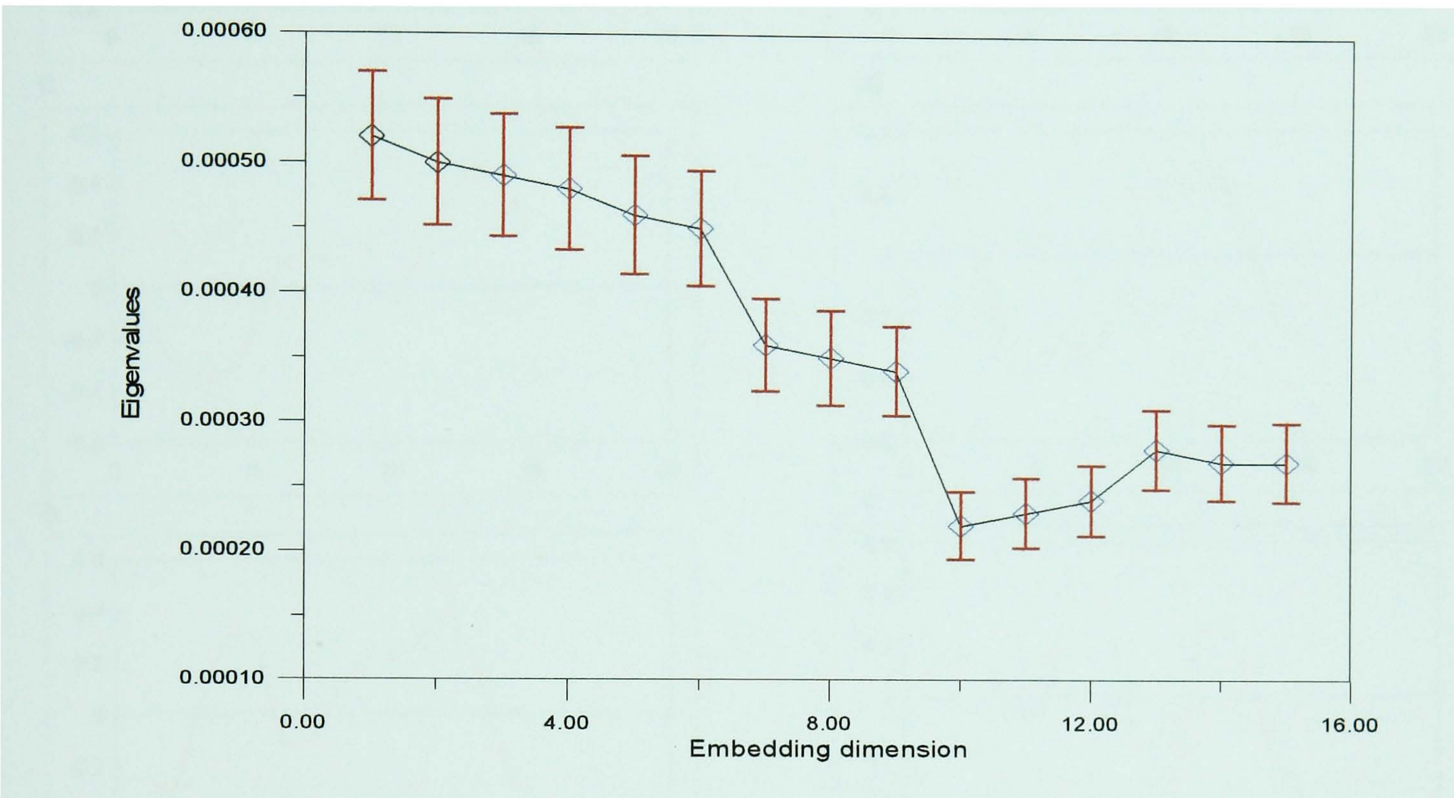

Figure 6.2. SVD analysis, ASE returns - Plot of the eigenvalues corresponding to diagonalization of a (15x15) covariance matrix, with 95% confidence limits calculated using the variance of the 1000 random realizations

To crosscheck this finding we also inspected the shape of the eigenvectors, as mentioned above. Figure 6.3a-j shows the plot of the first 10 eigenvectors produced by the SVD analysis with a (15x15) covariance matrix. It is clear that only the first 6 of them resemble plots of orthogonal polynomials, while for the rest the shape is increasingly disturbed.

These findings seem to support the existence of a deterministic part in the structure of our series. They also seem to support our correlation dimension estimate of $d \cong 6$ at least as the "statistical" dimension, i.e. an upper bound to the dimensionality of the deterministic part of the trajectory.

Singular vector Plots

(15x15) Covariance matrix

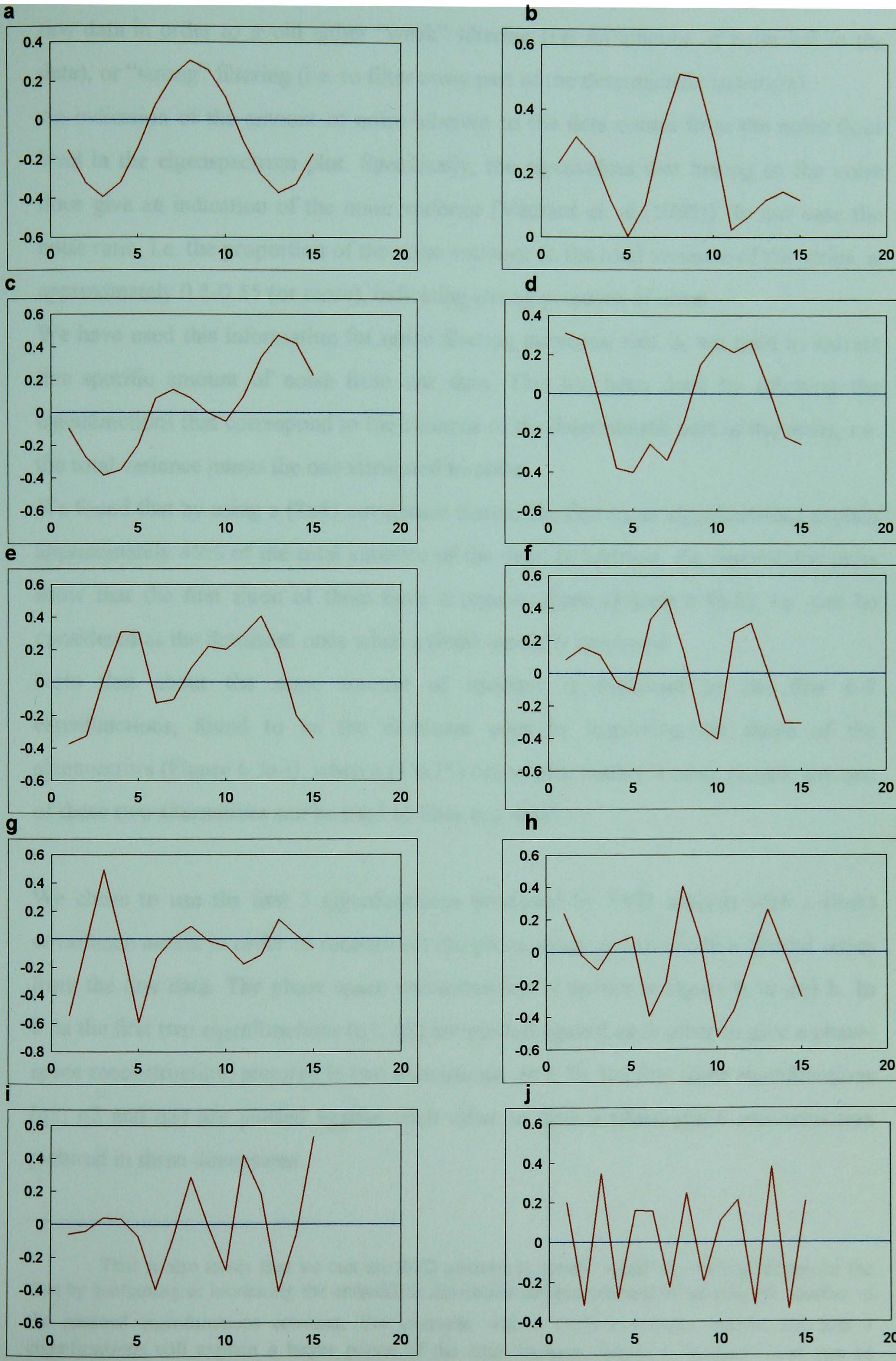

Figure 6.3 (a-j) The first 10 (out of 15) eigenvector plots for the ASE returns, after diagonalizing a (15x15) covar. matrix

With respect to the noise filtering process, the problem is to define the noise level of the raw data in order to avoid either "weak" filtering (i.e. an amount of noise left in the data), or "strong" filtering (i.e. to filter away part of the deterministic structure).

An indication of the amount of noise inherent to the data comes from the noise floor level in the eigenspectrum plot. Specifically, the eigenvalues that belong to the noise floor give an indication of the noise variance [Vautard et. al.(1992)]. In our case the noise ratio, i.e. the proportion of the noise variance to the total variance of the series, is approximately 0.5-0.55 (or more), indicating strong presence of noise.

We have used this information for noise filtering purposes, that is, we tried to extract this specific amount of noise from our data. This has been done by selecting the eigenfunctions that correspond to the variance of the deterministic part of the series, i.e. the total variance minus the one attributed to noise.

We found that by using a (8x8) covariance matrix, the first three eigenfunctions explain approximately 45% of the total variance of the data. In addition, the eigenvector plots show that the first three of them have a regular shape (Figure 6.4a-h), i.e. can be considered as the dominant ones when a (8x8) matrix is employed.

Note that about the same amount of variance is explained by the first 6-7 eigenfunctions, found to be the dominant ones by inspecting the shape of the eigenvectors (Figure 6.3a-j), when a (15x15) covariance matrix is used. Hence, any one of these two alternatives can be used to filter our data[2].

We chose to use the first 3 eigenfunctions produced by SVD analysis with a (8x8) covariance matrix in order to reconstruct the phase space and to create a filtered series from the raw data. The phase space reconstruction is shown in Figure 6.5a and b. In 6.5a the first two eigenfunctions (q1, q2) are plotted against each other to give a phase-space reconstruction, pictured in two dimensions. In 6.5b, the first three eigenfunctions (q1, q2 and q3) are plotted against each other to give a phase-space reconstruction pictured in three dimensions.

---

2 That is also to say that we can use SVD analysis to create "weak" or "strong" filters for the data by decreasing or increasing the embedding dimension respectively and by keeping the number of the retained eigenfunctions constant. For example, with a (5x5) covariance matrix, the first 3 eigenfunctions will explain a larger potion of the total variance, hence, a "weaker" filter can be constructed, i.e. a filter which will eliminate a smaller portion of the existing noise.

Singular vector Plots

(8x8) Covariance matrix

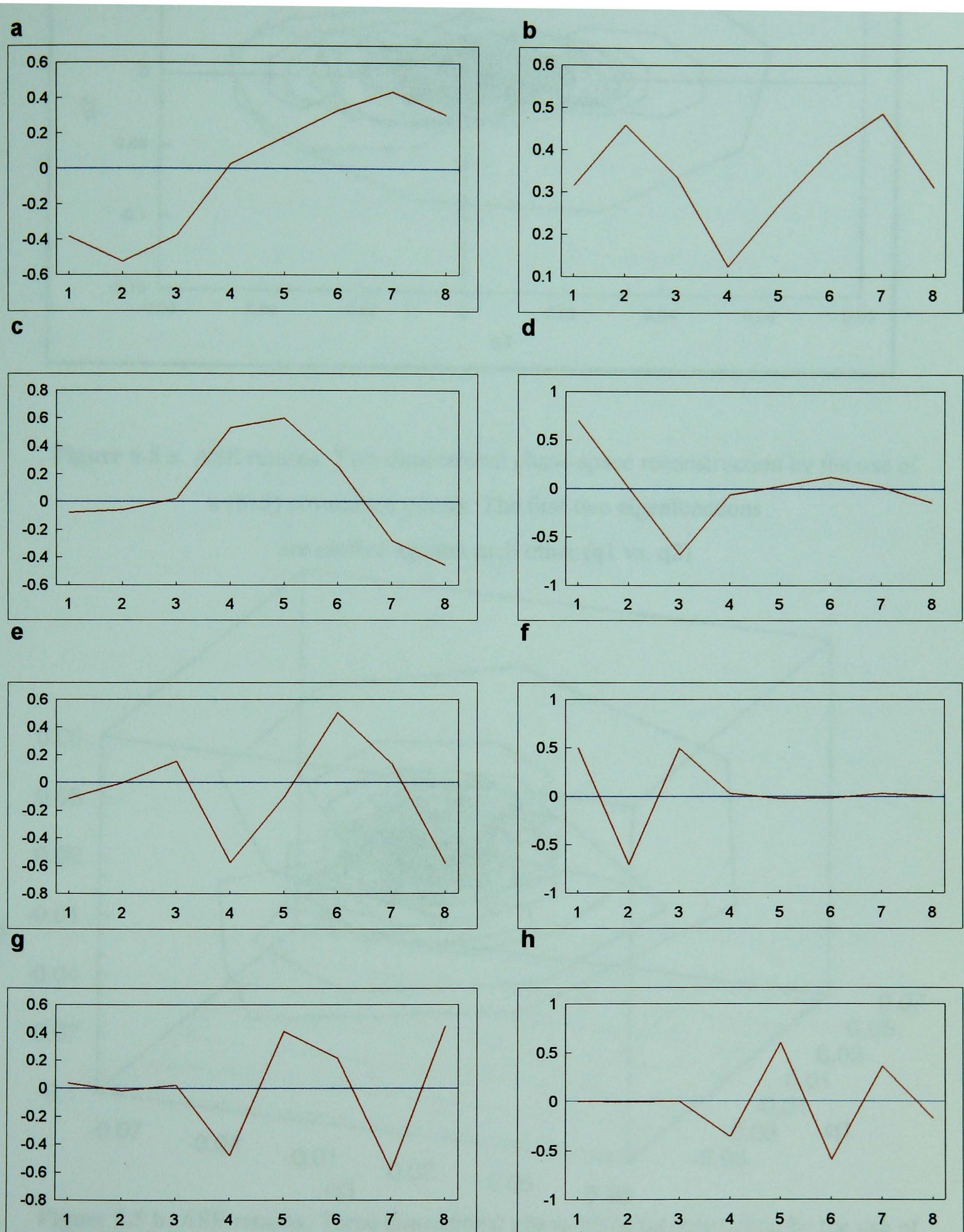

Figure 6.4 (a-h) 8 (out of 8) eigenvector plots for the ASE returns, after diagonalizing a (8x8) covariance matrix

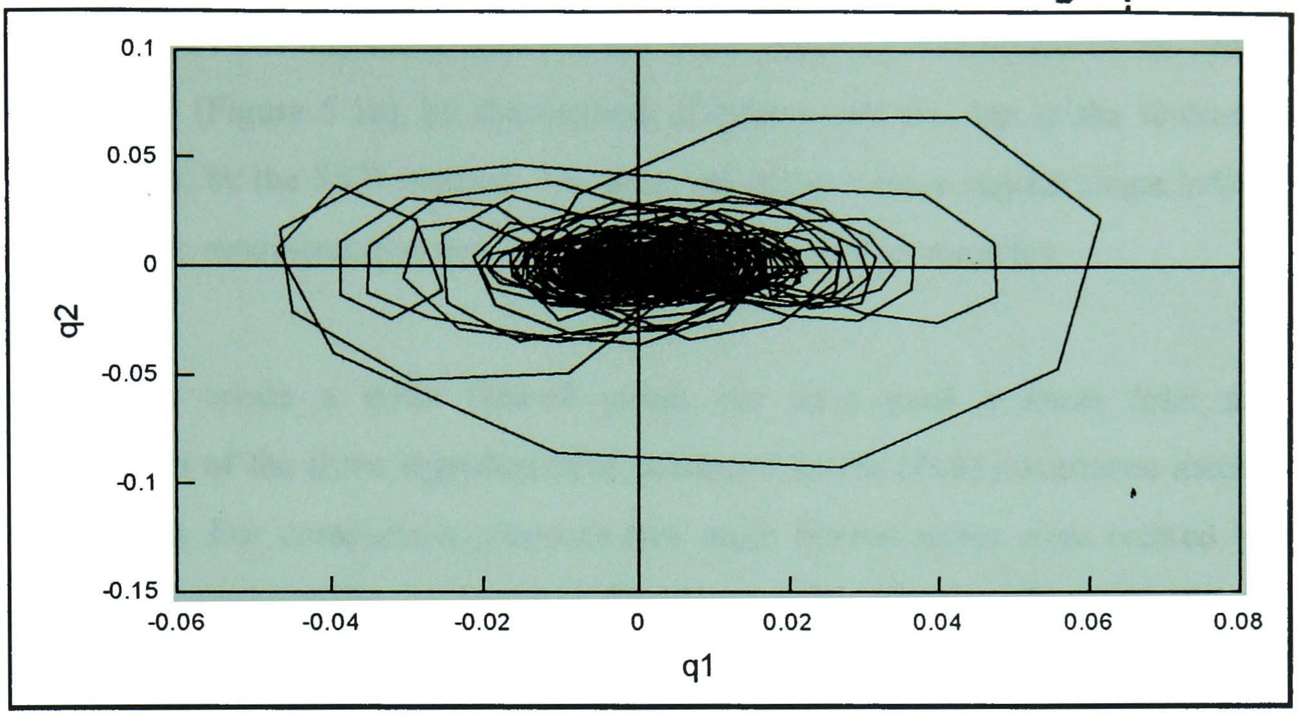

Figure 6.5 a. ASE returns: Two dimensional phase-space reconstruction by the use of a (8x8) covariance matrix. The first two eigenfunctions are plotted against each other (q1 vs. q2)

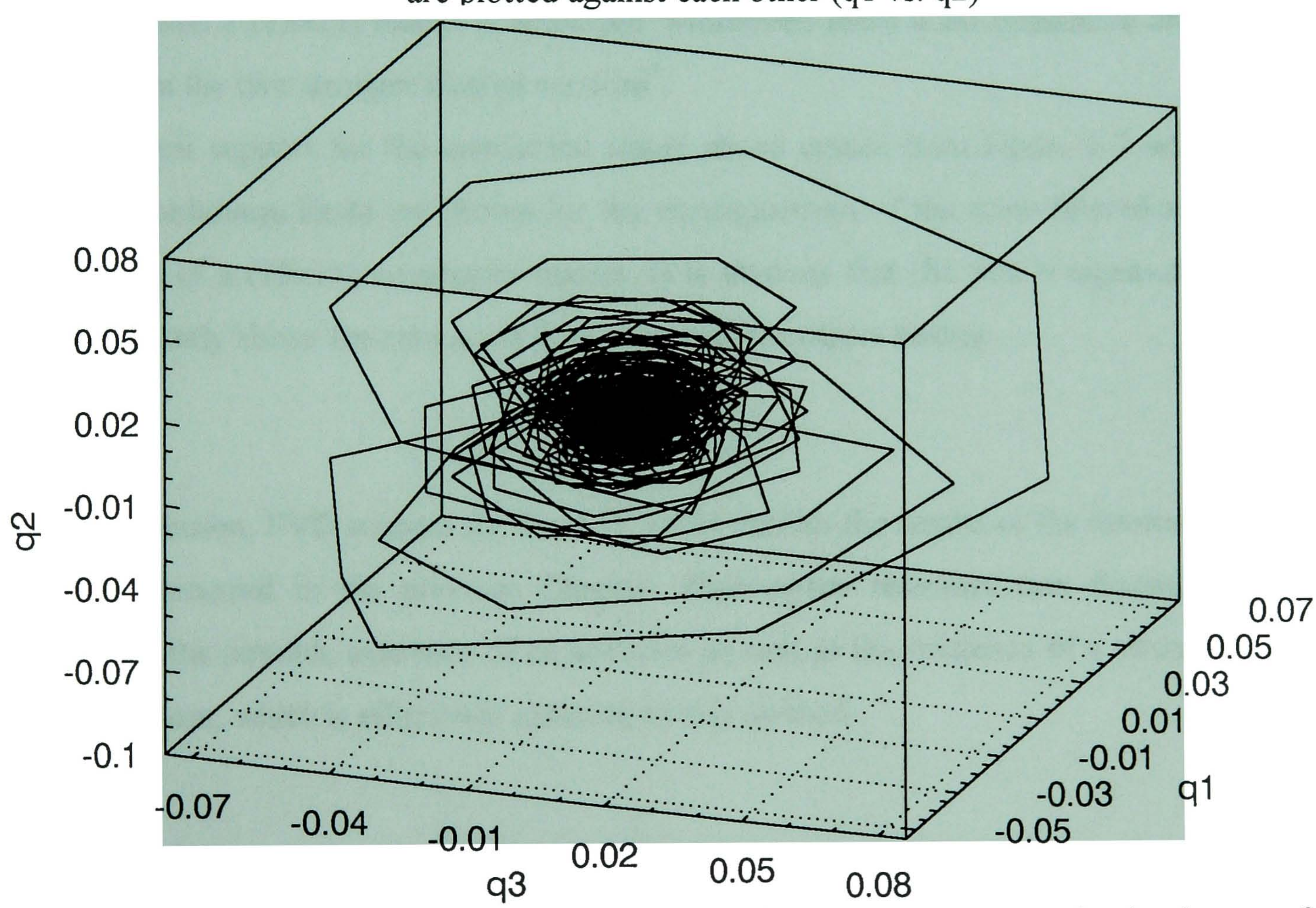

Figure 6.5 b. ASE returns: Three dimensional phase-space reconstruction by the use of a (8x8) covariance matrix. Here, the first three eigenfunctions are plotted against each other (q1 vs. q2 vs. q3)

There is an obvious difference between the phase space reconstruction of the raw data in Chapter 5 (Figure 5.1e), by the method of delays, and the one of the filtered data (Figure 6.5a), by the SVD method. The latter produces a more regular shape indicating a much better reconstruction and the possible existence of an attractor.

In order to create a noise filtered series, we have used a linear least square superposition of the three eigenfunctions produced by the (8x8) covariance matrix, to the raw data. For comparison purposes two more filtered series were created in the same way, corresponding to stronger filtering, i.e. by the use of a (10x10) and a (15x15) covariance matrix respectively[3].

The eigenspectrum of the noise-filtered series is shown in Figure 6.6 a,b,c for m = 8, 11 and 15 respectively. It is clearly shown that the noise floor has been eliminated in all cases and the number of the eigenvalues above zero is no more than 6, even when a (15x15) matrix is employed. Moreover, there is no qualitative difference between the two stronger filtered versions[4].

Statistical support for the conclusion stated above comes from Figure 6.7 where the 95% confidence limits are shown for the eigenspectrum of the noise filtered series by the use of a (15x15) covariance matrix. It is obvious that the first 6 eigenvalues are significantly above the zero noise floor where all the others belong.

In conclusion, SVD analysis for the ASE series verifies the results of the dimensionality tests discussed in the previous Chapter. Phase-space reconstruction through SVD reveals the possible existence of an attractor as well as the existence of a strong noisy component, which is effectively removed by this method.

---

3 The variance explained by the first 3 eigenfunctions of a (10x10) specification is 35% of the original variance while the corresponding Figure for the (15x15) specification is 22%.

4 We should notice, that this invariant behaviour of the eigenspectrum for $m>10$, might be interpreted as an urge for stronger filtering (e.g. by a (10x10) matrix) of the raw data.

ASE Noise Filtered Series

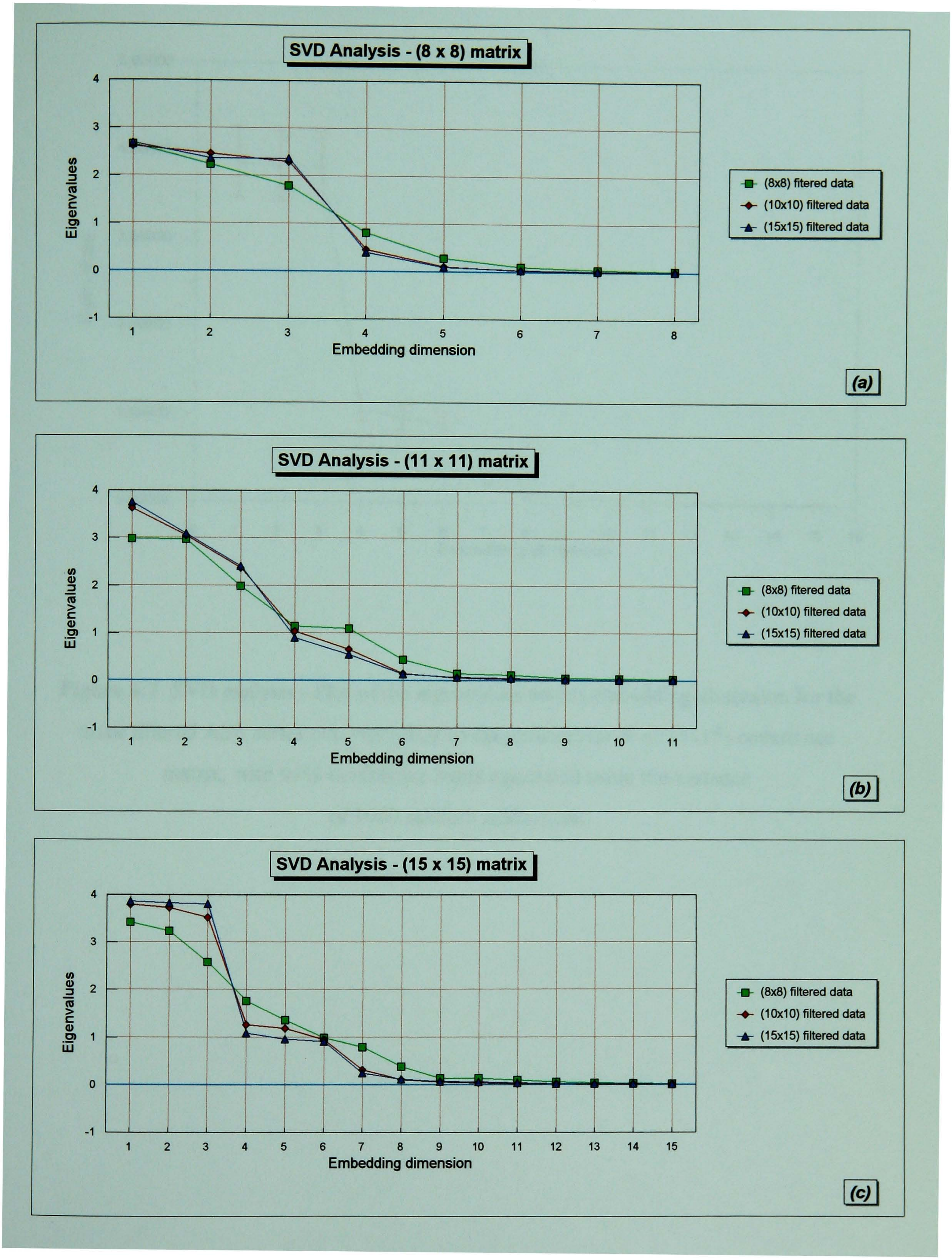

Figure 6.6(a-c) SVD analysis - Plot of the eigenvalues versus embedding dimension for the noise filtered ASE series. (a),(b) and (c) correspond to (8x8), (11x11) and (15x15) covariance matrices respectively. The three curves in each plot correspond to different levels (stronger or weaker) of noise filtering

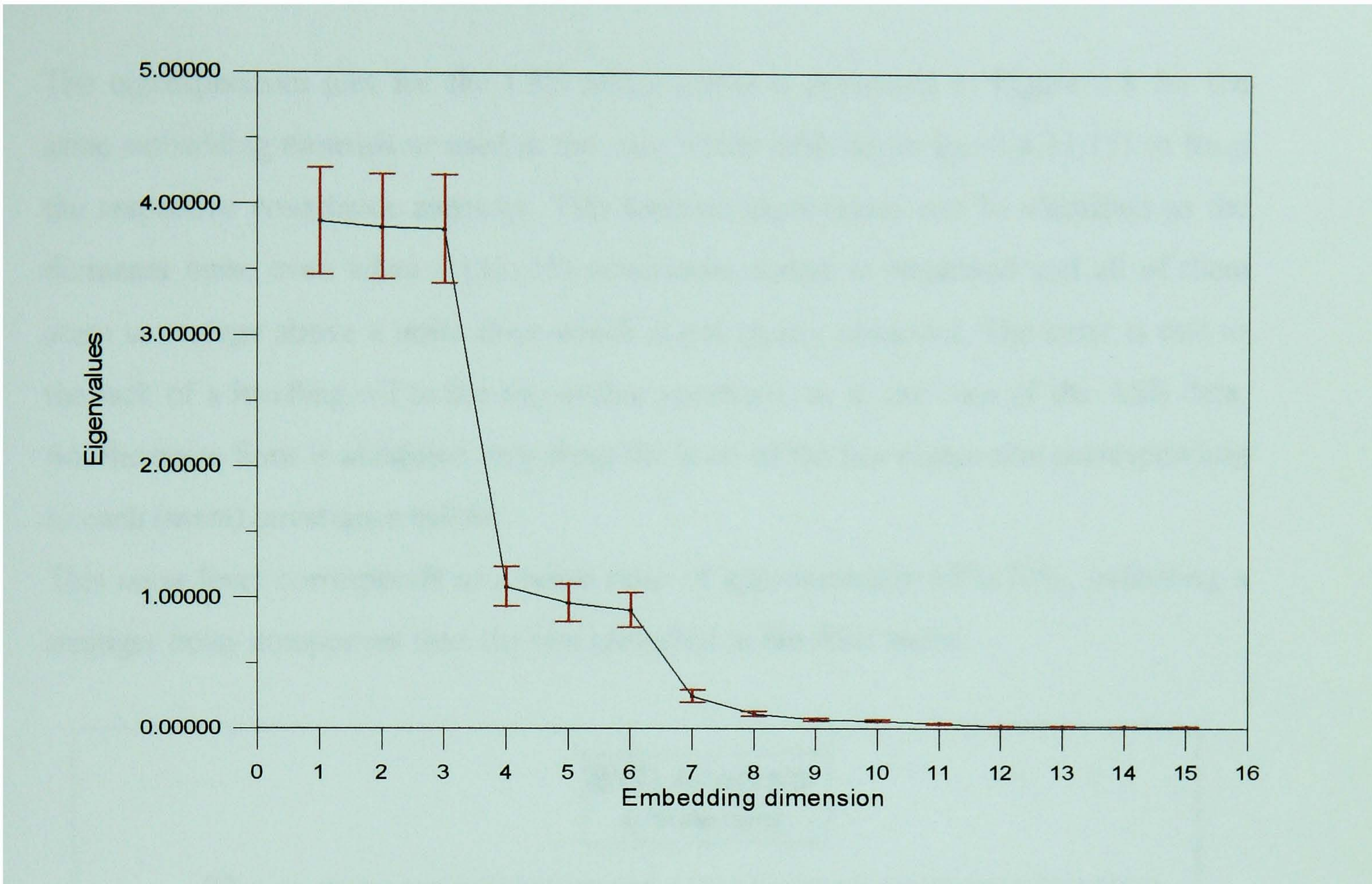

Figure 6.7. SVD analysis - Plot of the eigenvalues versus embedding dimension for the noise filtered ASE series corresponding to diagonalization of a (15x15) covariance matrix, with 95% confidence limits calculated using the variance of 1000 random realizations

6.1.2.2 The case of the LSE data

The eigenspectrum plot for the LSE return series is presented in Figure 6.8 for the same embedding dimensions used in the case of the ASE series (m=5,8,11,15) to form the respective covariance matrices. This time no eigenvalues can be identified as the dominant ones, even when a (15x15) covariance matrix is employed and all of them seem to emerge above a noise floor which is not clearly observed. The latter is due to the lack of a levelling-off in the eigenvalue spectrum, as in the case of the ASE data. So, the noise floor is identified only from the level of the last eigenvalue corresponding to each (mxm) covariance matrix.

This noise level corresponds to a noise ratio of approximately 65%-70%, indicating a stronger noisy component than the one identified in the ASE series.

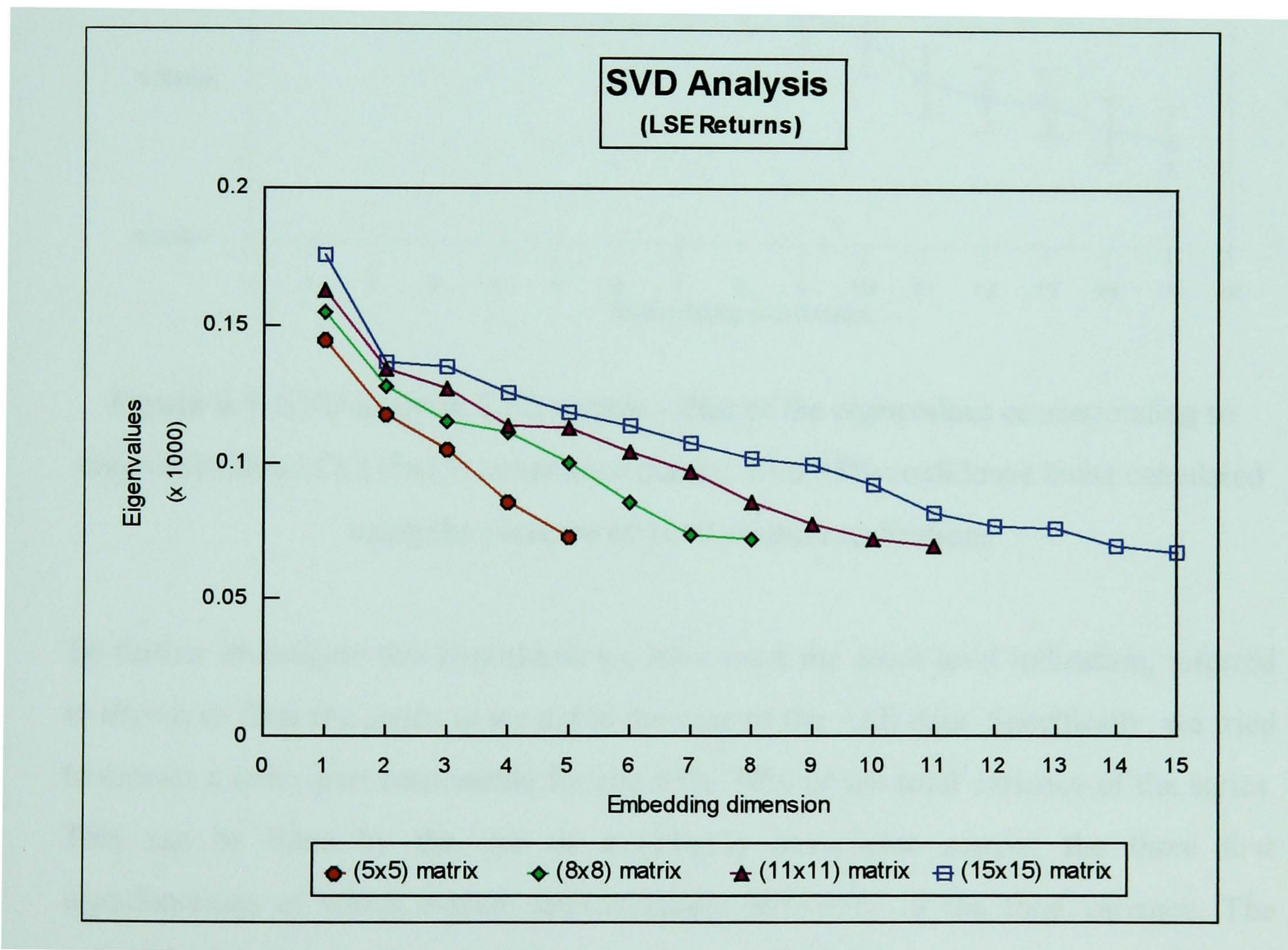

Figure 6.8. SVD analysis - Plot of the eigenvalues versus embedding dimension for the LSE returns. The four curves correspond to diagonalization of a (5x5), (8x8), (11x11) and (15x15) covariance matrix respectively

The same conclusion holds after inspecting the 95% error bars in Figure 6.9. With the exception of the first, the rest of the eigenvalues in the spectrum produced by the (15x15) covariance matrix cannot be significantly distinguished from each other and there are no clear indications to support the existence of a deterministic part in the structure of the series.

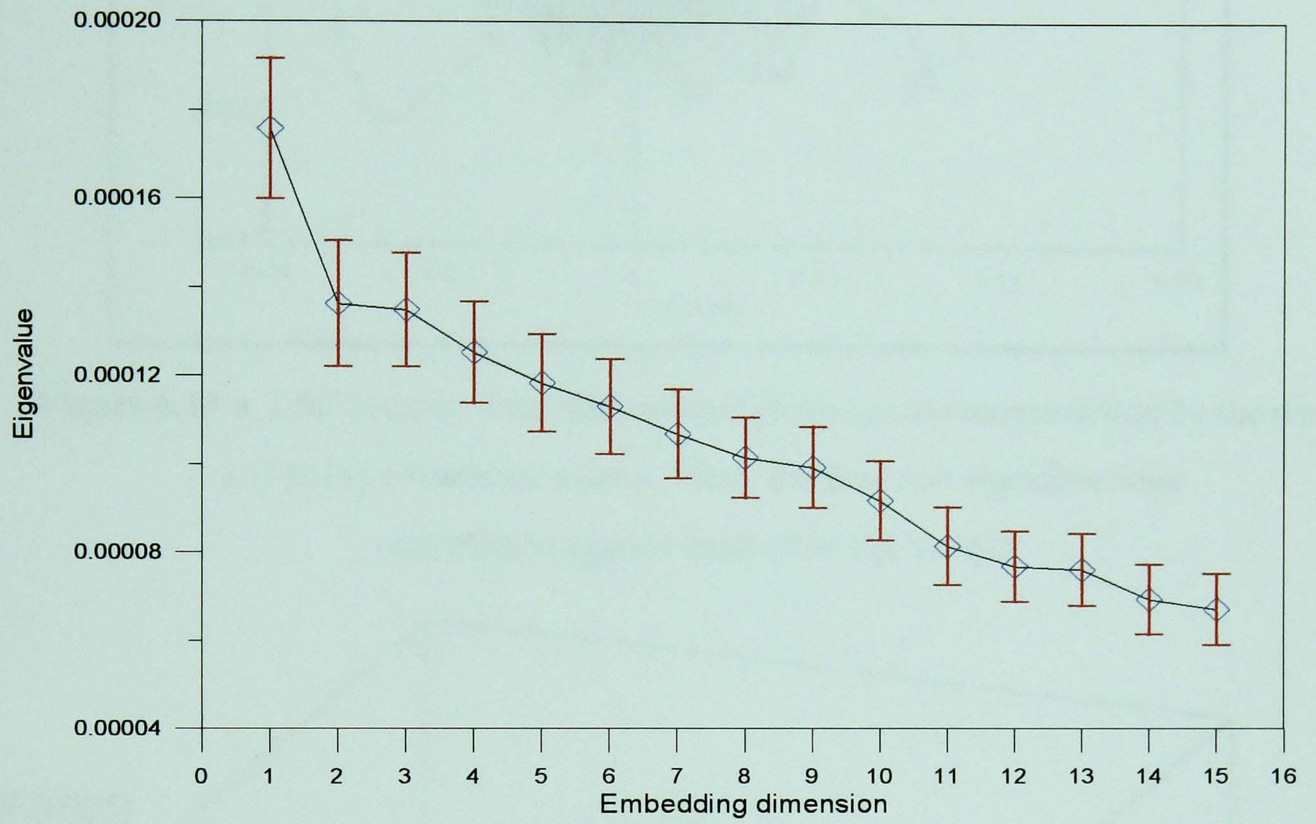

Figure 6.9. SVD analysis, LSE returns - Plot of the eigenvalues corresponding to diagonalization of a (15x15) covariance matrix, with 95% confidence limits calculated using the variance of 1000 random realizations

To further investigate this hypothesis we have used the noise level indication, referred to above, to filter the series as we did in the case of the ASE data. Specifically, we tried to extract a noisy part responsible for the 65%-70% of the total variance of the series. This can be done by the use of a (11x11) covariance matrix, the three first eigenfunctions of which explain approximately 30%-35% of the total variance. The noise filtered data has been produced again by a linear least square superposition of these three eigenfunctions to the raw LSE series. This time no additional filtering (stronger or weaker) was used, due to the high level of the noise component.

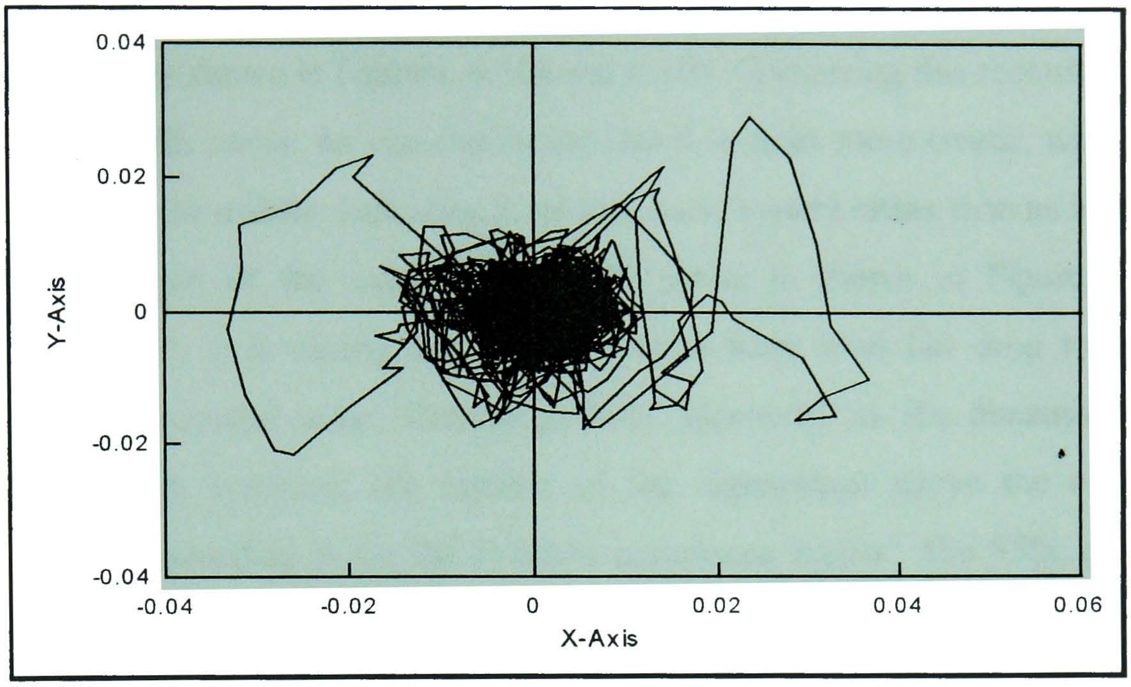

Figure 6.10 a. LSE returns: Two dimensional phase-space reconstruction by the use of a (11x11) covariance matrix. Here, the first two eigenfunctions are plotted against each other (q1 vs. q2)

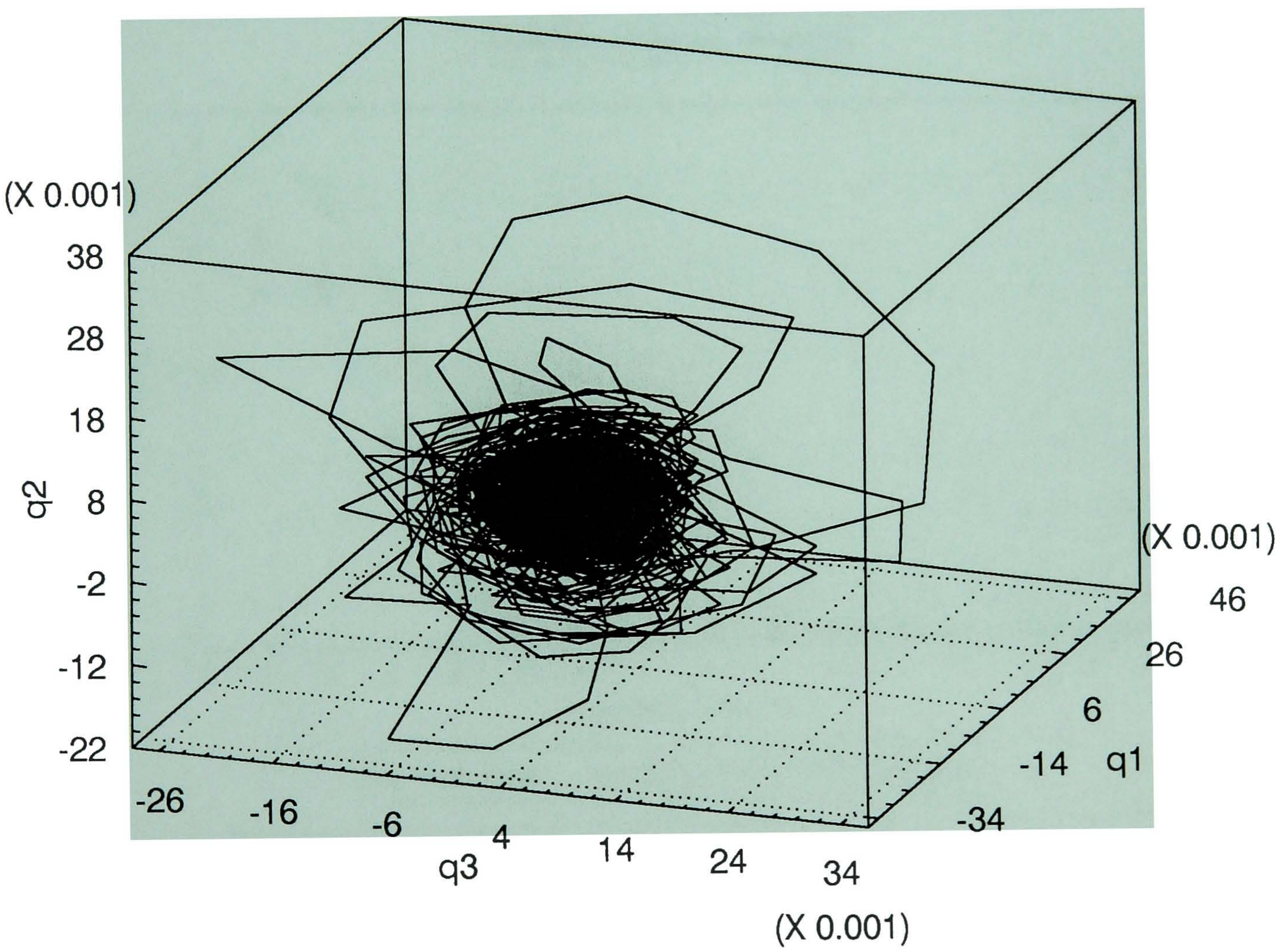

Figure 6.10 b. LSE returns: Three dimensional phase-space reconstruction by the use of a (11x11) covariance matrix. Here, the first three eigenfunctions are plotted against each other (q1 vs. q2 vs. q3)

The same eigenfunctions have been used to reconstruct the phase space of the system and the results are shown in Figures. 6.10a and 6.10b. Comparing this reconstruction to the one of the ASE series, we can clearly see that it is much more erratic, with a heavy concentration in the middle, indicating a random noisy system rather than an attractor. The eigenspectrum of the noise-filtered LSE series is shown in Figure 6.11 for m = 8,11 and 15. It is clearly seen that the noise floor level has drop to zero, an indication of successful noise filtering process. However, as the dimension of the covariance matrix increases, the number of the eigenvalues above the noise floor increases, too, exceeding 8 for the (15x15) covariance matrix. The 95% confidence intervals in Figure 6.12 once again verify this conclusion. This finding is in direct agreement with our correlation dimension estimations for the LSE series, where a higher than 8 and non-saturating dimension was reported.

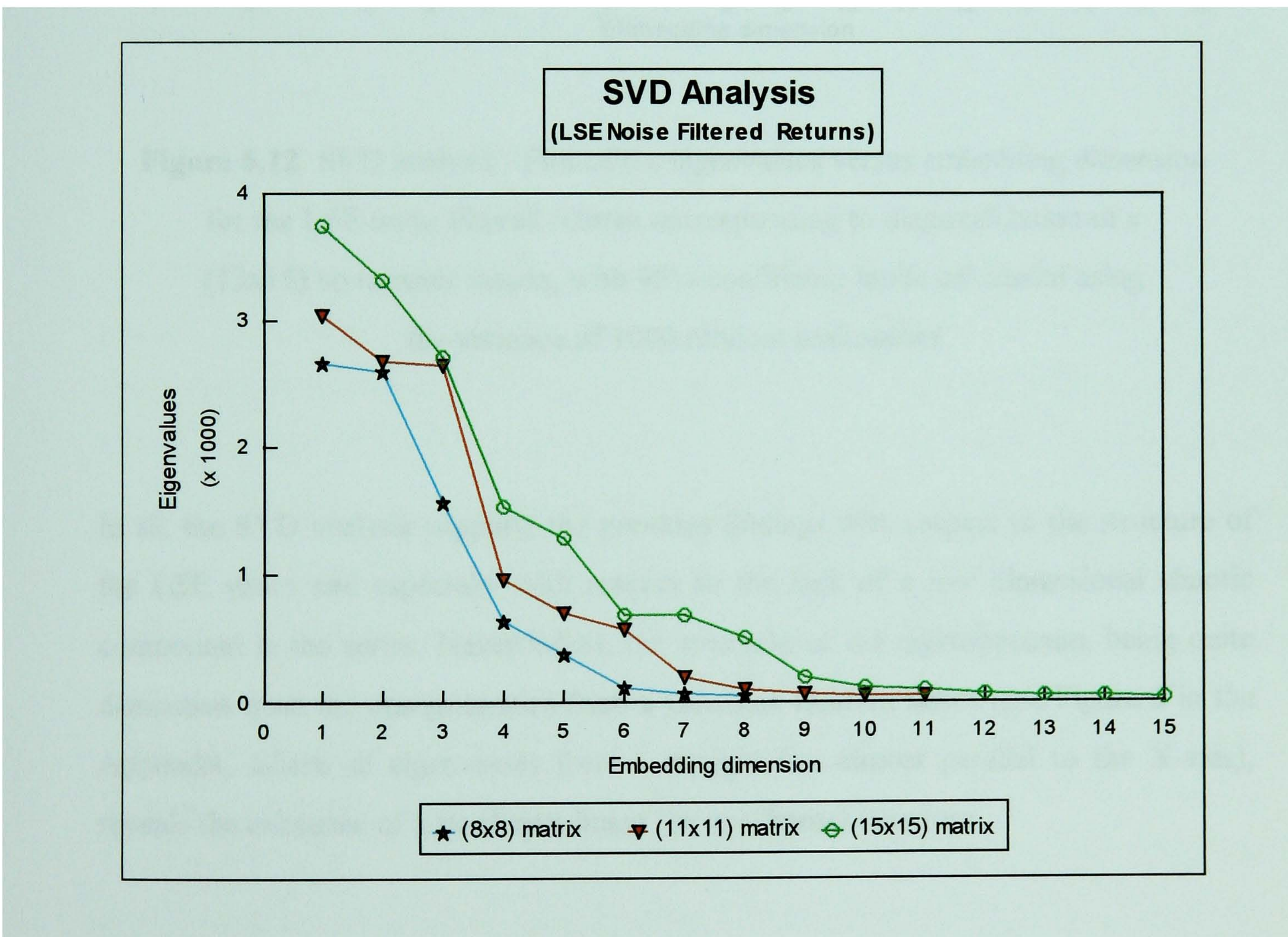

Figure 6.11. SVD analysis - Plot of the eigenvalues versus embedding dimension for the LSE noise filtered returns. The three curves correspond to diagonalization of a (8x8), (11x11) and (15x15) covariance matrix respectively .

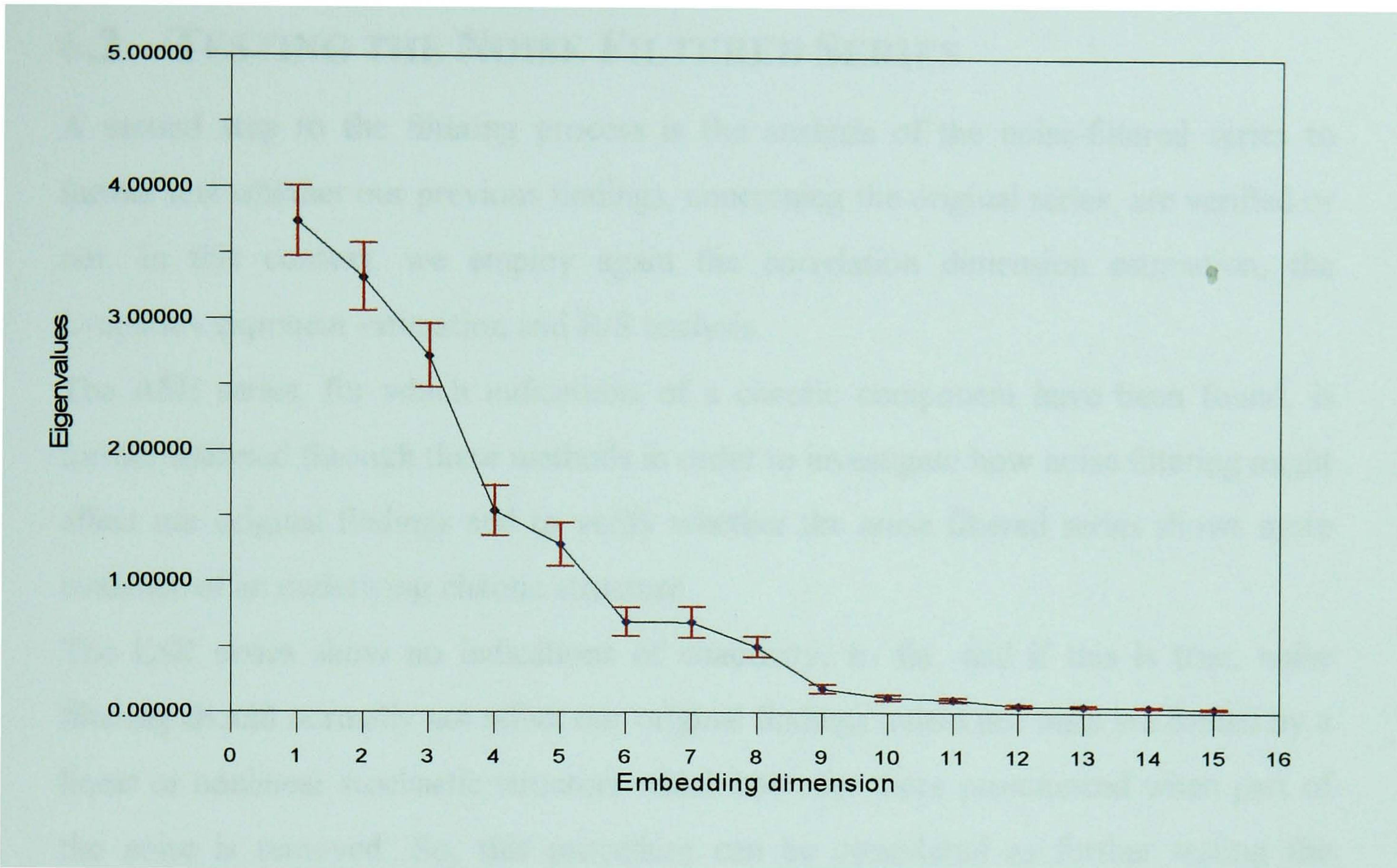

Figure 6.12. SVD analysis - Plot of the eigenvalues versus embedding dimension for the LSE noise filtered returns corresponding to diagonalization of a (15x15) covariance matrix, with 95% confidence limits calculated using the variance of 1000 random realizations

In all, the SVD analysis supports the previous findings with respect to the structure of the LSE series and especially with respect to the lack of a low dimensional chaotic component in the series. Nevertheless, the structure of the eigenspectrum, being quite distinctive from the one generated from a Gaussian random series (see Figure 2 in the Appendix, where all eigenvelues form a straight line almost parallel to the X-axis), reveals the existence of a stochastic linear (or non-linear) structure.

6.2 TESTING THE NOISE FILTERED SERIES

A second step to the filtering process is the analysis of the noise-filtered series to further test whether our previous findings, concerning the original series, are verified or not. In this context, we employ again the correlation dimension estimation, the Lyapunov exponent estimation and R/S analysis.

The ASE series, for which indications of a chaotic component have been found, is further analysed through these methods in order to investigate how noise filtering might affect our original findings and to verify whether the noise filtered series shows more evidence of an underlying chaotic structure.

The LSE series show no indications of chaoticity, so far, and if this is true, noise filtering should normally not affect our original findings unless our tests are fooled by a linear or nonlinear stochastic structure which becomes more pronounced when part of the noise is removed. So, this procedure can be considered as further testing the reliability of our methods in detecting chaotic components.

In terms of basic statistical properties, the two series, as expected from the filtering process, show smaller variance, skewness and kurtosis in comparison with the original series. The differences are more pronounced in the time and frequency domain, as reflected in their autocorrelogram and power spectrum respectively. This is shown in Figure 6.13, where, for comparison purposes, the autocorrelogram and the power spectrum of four different series are presented, including our two filtered series, a fractional noise series (H=0,65) and a chaotic series (Lorenz attractor).

It can be shown that both our series exhibit an autocorrelation signature with long and slowly decaying autocorrelations, met also in many chaotic processes (e.g. Lorenz attractor, Ikeda map, Henon attractor) as well as in random fractal processes. The same is true for the power spectrum, which is very similar in our two series and shows, in its high frequency part, a structure similar to that of fractional ($1/f^a$) noise. It is obvious from the Figure above that, as already discussed in previous Chapters, no conclusions can be drawn concerning the structure of our series based on standard analysis in the time and frequency domains.

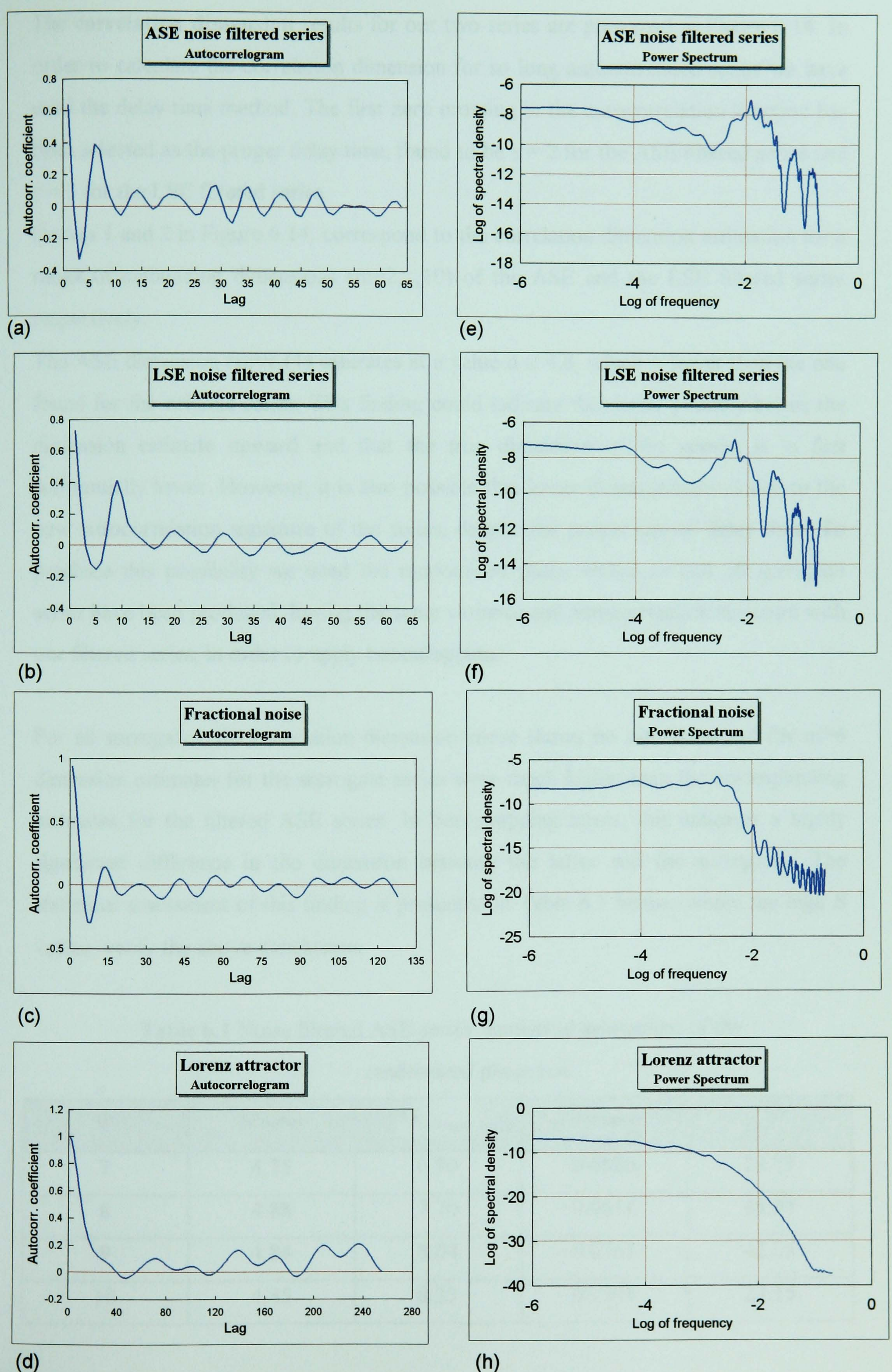

Figure 6.13 (a-d) : Autocorrelogram of the ASE and LSE noise filtered series, a Fractional noise series and Lorenz attractor (e-h): Power Spectrum of the same series

The **correlation dimension** results for our two series are presented in Figure 6.14. In order to calculate the correlation dimension for so long autocorrelated series we have used the delay time method. The first zero crossing in the autocorrelation function has been selected as the proper delay time, found to be $\tau = 2$ for the ASE filtered series and $\tau = 3$ for the LSE filtered series.

Curves 1 and 2 in Figure 6.14, correspond to the correlation dimension estimation for a range of embedding dimensions (m=2,..,10) of the ASE and the LSE filtered series respectively.

The ASE dimension curve (1) saturates at a value $d \cong 4.8$, which is lower than the one found for the original series. This finding could indicate that noise possibly biases the dimension estimate upward and that the true dimension of the system is in fact substantially lower. However, it is also possible that lower dimensionality is due to the new autocorrelation signature of the series, despite the proper use of delay time. To preclude this possibility we used the randomized phase technique and 50 surrogate series have been produced, having the same variance and autocorrelation structure with our filtered series, in order to apply bootstrapping.

For all surrogates, the correlation dimension curve shows no saturation and for m>6 dimension estimates for the surrogate series were much higher than the corresponding estimates for the filtered ASE series. In bootstrapping terms, this indicates a highly significant difference in the dimension between the latter and the surrogates. The statistical assessment of this finding is presented in Table 6.1 below, where the high S values, verify the above conclusion.

Table 6.1 Noise filtered ASE series: Statistical assessment of the randomized phase test

[illegible]	$Q_{original}$	$\overline{Q}_{surrogate}$	$S_{surrogate}$	[illegible]
7	4.75	6.70	0.0820	23.77
8	4.88	7.70	0.0614	45.93
9	4.84	8.04	0.0761	42.05
10	4.85	8.23	0.1245	27.15

The mean correlation dimension estimate of the 50 surrogates for each m is represented by curve (3) in Figure 6.14, where the lack of saturation and the higher dimension estimates of the surrogates can be clearly observed.

The LSE dimension curve (2) in Figure 6.14 shows no saturation and the correlation dimension estimates for each m are very similar to the ones measured in the original series. This is a strong indication in favor of our findings so far, but also a verification of the ability of the correlation dimension method to distinguish between lower (possibly chaotic) and higher dimensional (stochastic) processes.

To cross check the reliability of this result, we used again the "independent realisations" method, and a filtered series has been produced for the "short" LSE series, being directly comparable to the ASE series. In terms of SVD analysis, the "short" data series exhibits an eigenspectrum structure identical to that of the long series, but the noise level is a little higher (70%-75%). To filter this series, we used a linear least square superposition of the first three eigenfunctions produced by a (13x13) covariance matrix to the "short" LSE series. It was found (by a trial and error procedure with covariance matrices of varying dimension) that this filtering process removes approximately 75% of noise. Then the correlation dimension of the "short" filtered series was estimated and the results are presented in Figure 6.14, (curve 4). As we can see, the dimension estimate is of the same magnitude to that of the longer original series, indicating once again that the dimension estimate of the LSE filtered series is not sensitive to the data length or the period tested. Indirectly, this means that our results for the noise filtered ASE series are reliable and not due to short data length bias.

With respect to **the R/S analysis**, because of the amount of autocorrelation in the filtered data which makes removal of short-term dependence difficult and impractical, we chose to apply the modified R/S statistic, which accounts for short-term dependence. Before that, we estimated the V-statistic curve for both data sets to see whether its structure is preserved after noise filtering. Recall that short-term dependence moves the V-statistic slope upward and leads to biased Hurst estimate, but does not affect the cyclycity of a series. The V-statistic plots for the ASE and the LSE filtered series are presented in Figure 6.15(a)-6.15(b), respectively. As we can see from these plots the ASE filtered series

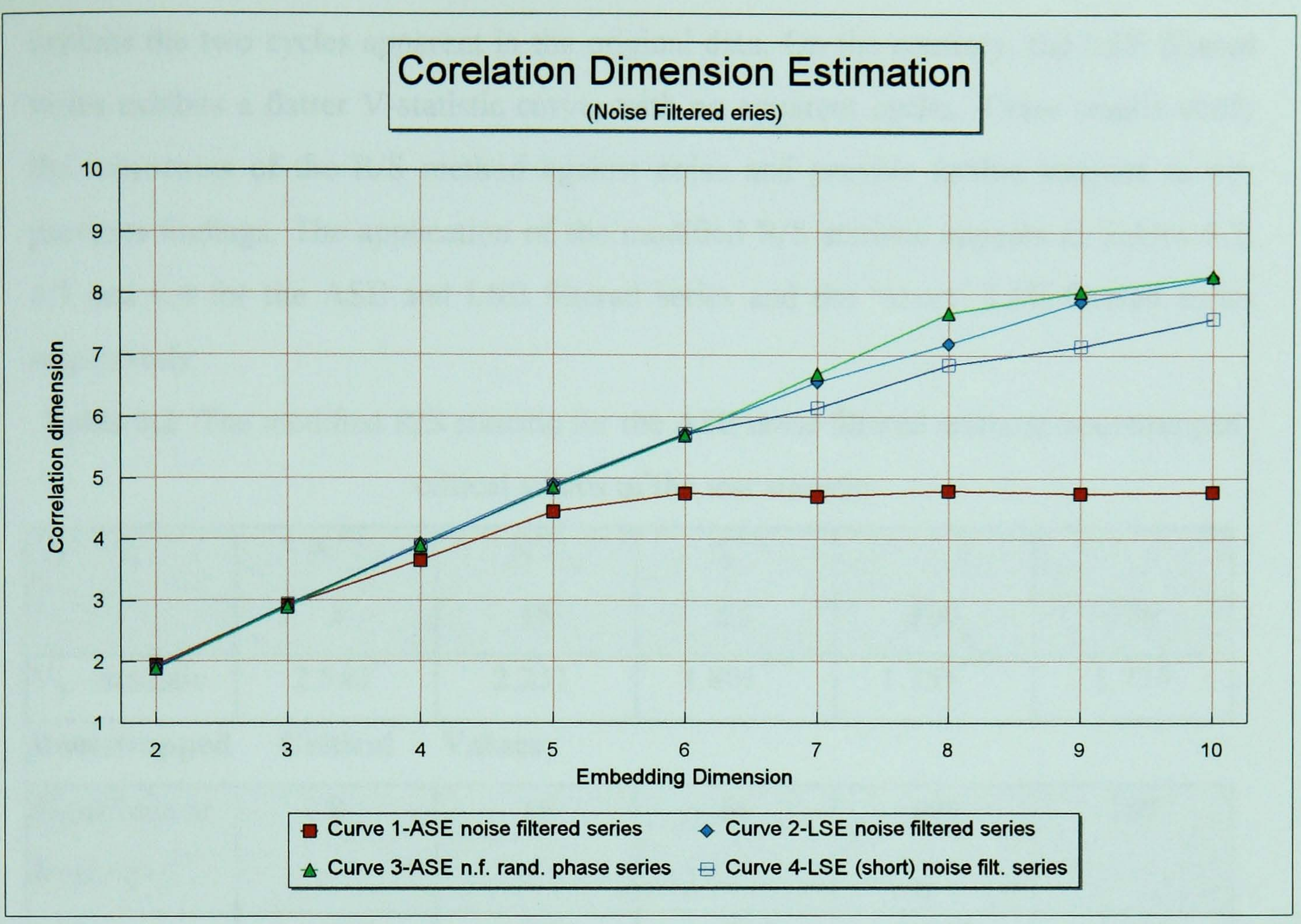

Figure 6.14 Plot of the correlation dimension estimate (d) versus embedding dimension (m) for the noise filtered ASE series and its randomized phase surrogate, the noise filtered LSE series and the "short" noise filtered LSE series

R/S Analysis - V Statistic plot

ASE filtered series

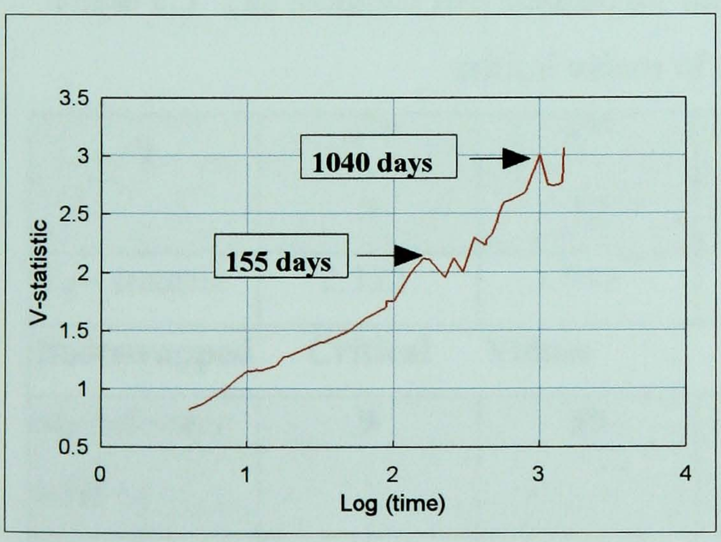

Figure 6.15 (a) V-statistic plot of the ASE noise filtered series. Cycles are indicated by arrows and the V-statistic curve is similar to that of the original ASE series

LSE filtered series

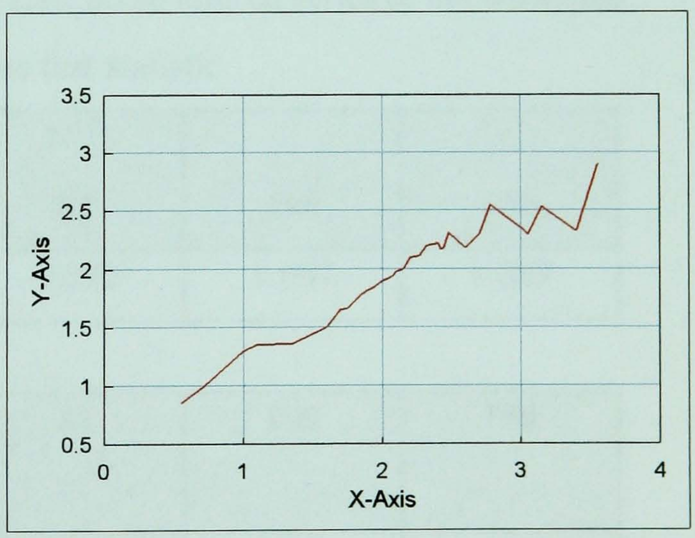

Figure 6.15 (b) V-statistic plot of the LSE noise filtered series. No cycle is discernible, and the V-statistic curve is similar to that of the original series.

exhibits the two cycles apparent in the original data. On the contrary, the LSE filtered series exhibits a flatter V-statistic curve, with no apparent cycles. These results verify the robustness of the R/S method against noise and provide further support to our previous findings. The application of the modified R/S statistic appears in Tables 6.2, 6.3 and 6.4 for the ASE and LSE filtered series and the "short" LSE filtered series respectively.

Table 6.2 The modified R/S statistic for the ASE noise filtered series & bootstrapped critical values of the test statistic

q	**$N^{1/4}$ 8**	**$N^{1/3}$ 15**	**$N^{1/2}$ 56**	**100**	**150**
V_q – statistic	2.540^{*}	2.231^{*}	1.801^{***}	1.755^{***}	1.730
Bootstrapped Critical Values					
Significance level / q	**8**	**15**	**56**	**100**	**150**
1.0%	1.885	1.875	1.869	1.807	1.779
2.5%	1.796	1.784	1.765	1.702	1.697
5.0%	1.695	1.670	1.656	1.629	1.625

* Significance at one-tail 1.0% level according to the asymptotic critical values
** Significance at one-tail 2.5% level according to the asymptotic critical values
*** Significance at one-tail 5.0% level according to the asymptotic critical values

Table 6.3 The modified R/S statistic for the LSE noise filtered series & bootstrapped critical values of the test statistic

q	**$N^{1/4}$ 9**	**$N^{1/3}$ 19**	**$N^{1/2}$ 82**	**100**	**150**
V_q – statistic	2.332^{*}	1.942^{**}	1.690	1.667	1.689
Bootstrapped Critical Values					
Significance level / q	**9**	**19**	**82**	**100**	**150**
1.0%	1.913	1.887	1.875	1.871	1.865
2.5%	1.835	1.809	1.800	1.790	1.762
5.0%	1.750	1.715	1.709	1.722	1.708

* Significance at one-tail 1.0% level according to the asymptotic critical values
** Significance at one-tail 2.5% level according to the asymptotic critical values

Table 6.4 The modified R/S statistic for the "short" LSE noise filtered series & bootstrapped critical values of the test statistic

q	**$N^{1/4}$** **8**	**$N^{1/3}$** **15**	**$N^{1/2}$** **58**	**100**	**150**
V_q – statistic	1.663	1.407	1.288	1.341	1.423
Bootstrapped	**Critical**	**Values**			
Significance level / q	**8**	**15**	**58**	**100**	**150**
1.0%	1.977	1.925	1.853	1.868	1.792
2.5%	1.867	1.797	1.740	1.760	1.711
5.0%	1.739	1.710	1.665	1.682	1.637

The results presented above show that the short-term dependence null is rejected for the ASE series but not for the two LSE series using either the asymptotic or the bootstrapped critical values of the test statistic. Hence, as in the case of the ASE original series, the R/S test shows that long-term dependence and fractality, compatible with a chaotic explanation, cannot be ruled out for the ASE filtered series, either. On the contrary, the two filtered LSE series show no indication of fractality and short-term dependence seems to be their prominent characteristic.

The **Lyapunov exponent** of the ASE filtered series, was found to be positive and of the same magnitude as in the original series. Since R/S analysis of the filtered ASE series shows that cycles remain unchanged, we have used the same parameters and the $Q = m\tau$ rule, described in the previous Chapter, to calculate the Largest Lyapunov Exponent of the series. The results are presented in Table 6.5 and, as we can see, are directly comparable to the results of the original series in Table 5.5 of Chapter 5.
However, as in the case of the original series, the exponent of the LSE series was once again found positive and for both series the exponent was found to be increasing with the embedding dimension, revealing the difficulty of this measure in distinguishing between low and high dimensional processes, even in the absence of noise.

Table 6.5 LLE estimation for the ASE noise-filtered returns,

(t_s = 15 and m = 6, 9, 13)

[illegible]	[illegible]	[illegible]
6	26	0.017
9	17	0.010
13	12	0.006

The LLE estimates for the noise filtered ASE and LSE series for the same τ, m and t_s parameters as in the case of the original series (Table 5.7 in Chapter 5), are presented in Table 6.6 below. Notice that Lyapunov estimates for both the noise-filtered series remain practically identical to their counterparts for the original series.

Table 6.6 LLE for the ASE and LSE noise filtered series

m	t_s	L_1 (n.f. ASE series)	L_1 (n.f. LSE series)
6	5	0.0255	0.0525
9	5	0.0182	0.0368
13	5	0.0105	0.0111
6	10	0.0190	0.0365
9	10	0.0138	0.0244
13	10	0.0089	0.0080
6	15	0.0144	0.0287
9	15	0.0102	0.0201
13	15	0.0068	0.0077

Recapitulating, SVD analysis is a very useful tool in removing noise from a series. This ability, combined with further analysis of the filtered series, can help in distinguishing between lower and higher dimensional processes which look identical under standard analysis in the time and frequency domain.

Our results here provide further support towards a possible chaotic structure strongly mixed with noise in the ASE returns and are in line with the results from the preceding analysis. Noise reduction lowers the correlation dimension of the series, thus revealing the possibility of a lower dimensional structure. On the contrary, the LSE series do not show any indication of such a structure and are definitely high dimensional and short term dependent. Nonlinearity of the latter series is not ruled out, but it seems to be of a purely stochastic nature.

Chapter 7

NONLINEAR PREDICTION METHODS AS DETECTION TOOLS FOR CHAOTIC DYNAMICS

7.1 INTRODUCTION

In Chapters 7 and 8 we shall focus on forecasting issues. In this Chapter we shall show how non-linear forecasting methods can be used as tools for distinguishing between chaotic and stochastic processes as well as for specifying a reliable upper bound to the dimensionality of the putative attractor. In the empirical context, these methods will be used to further test our findings concerning the ASE data for which the preceding analysis does not rule out a chaotic explanation[1].

Traditionally, forecasting theory views time series as realisations of random processes, that is, processes involving many independent and irreducible degrees of freedom. However, recent developments have shown that "random looking" series may occur by chaotic systems with a few degrees of freedom which are short-term predictable with accuracy once the nonlinear model which generates the chaotic behaviour is known. In this case the problem is how to describe the asymptotic behaviour of iterates of a dynamical system given the nonlinear map which gives rise to them.
Yet, in most cases we are dealing with the opposite problem, i.e. given a sequence of iterates find the nonlinear map that generates them. If such a map can be found, short-term forecasting is possible and the process can be characterised as chaotic. With the exception of a few simple chaotic processes this is a rather difficult task especially when dealing with real data.

In the recent literature different non-linear modelling and forecasting techniques have been developed and can be used to investigate whether the irregularity in a time series

---

1 In all the different applications that have been used so far, the LSE data show no indications of low deterministic dynamics whatsoever, so no further testing of its dynamics through forecasting techniques was deemed necessary. However, our ability to forecast LSE returns will be tested in the next Chapter.

is due to low dimensional chaotic dynamics as opposed to high dimensional dynamics or stochasticity.

Assuming that the underlying dynamics of a system evolving in an m-dimensional space can be written as a map of the form:

$$\mathbf{x}(t+T) = f_T(\mathbf{x}(t)), \qquad (7.1)$$

where $\mathbf{x}(t)$ is the current state of the system and $\mathbf{x}(t+T)$ is its future state after time interval T, the problem is to approximate the unknown function f (both f and $\mathbf{x}$ are m-dimensional vectors), which should be non-linear if the system is chaotic, in order to be able to predict $\mathbf{x}(t+T)$. The correct approximation is usually tested by out-of-sample short-term forecasts.

In practice, given a scalar time series $\{x_t\}_{t=1,\ldots,N}$, the first step is to embed it in a state space and approach the problem as a «fitting» process of the form:

$$\mathbf{x}_{t+T} = f_T(x_t, x_{(t-1)\tau}, \ldots.., x_{t-(m-1)\tau}), \qquad (7.2)$$

where m = embedding dimension and τ = delay time of the delay vector $\mathbf{x}_t$. The approximation of the f map in this case is tested by out-of-sample short-term forecasts. Different techniques of approximating f are available and can be divided into two broad categories, namely global and local ones.

Global techniques often use polynomials as predictors, where the term "global" refers to the use of the whole data set for the fitting process. The simplest example of a global technique is the Autoregressive (AR) model [Priestley (1981)], a linear model using a first order polynomial approximator. In the case of complex chaotic systems global polynomial fits are impractical, since they usually involve a very large number of free parameters and non-linear expressions for f [Tsonis (1992), Casdagli (1989)]. A way to overcome this problem is to use rational predictors, i.e. the ratio of two polynomials [Casdagli (1989)], or multi-layer perceptron (MLP) neural networks [Lapedes and Farber (1987), Weigent et. al. (1990,1992)]. Alternatively, semi-local methods can be used, such as the radial basis functions (RBF), which are global interpolation techniques with good localisation properties [Carlin (1991)].

Local techniques construct local predictors that use only the states near the current state to make predictions, i.e. the nearest neighbours to the current state vector.

Geometrically, if a dynamical system is assumed to follow a trajectory s(t), local portions of this curve in the past are employed and predictions are based on the behaviour of the system immediately after these past events.

Nearest-neighbour techniques can be further divided to first-order approximations or "analogs", which use only one neighbour to construct the local predictor [Kennel and Isabelle (1992)], and second-order approximations which use more than one neighbours for the same purpose. Simplicity and computational efficiency are the main advantages of these techniques, the main representatives of which are the piecewise linear polynomial approximation [Farmer and Sidorowich (1987), Casdagli, (1991)] and the simplex method [Sugihara and May (1990)].

The nearest neighbour methods have been applied for prediction purposes by several researchers [Pawelzik and Schuster (1991), Townshent (1992), Hunter (1992), Cortini and Barton (1993)]. However, these techniques can be considered as a useful supplement to the testing framework used to distinguish between low-dimensional systems and stochastic ones, since in some cases they give less ambiguous results than traditional chaotic methods such as dimension calculations [Casdagli (1991)]. Combining the results of these methods is expected to further enhance our ability to tell whether the series under scrutiny exhibit indeed chaotic characteristics.

Sugihara and May (1990) also suggest that prediction methods, like the ones described in the next section, can be a more reliable tool in specifying an upper bound to the dimensionality of an attractor than the usual correlation dimension procedure. This is so, because prediction methods do not suffer from data limitation problems and are able to exploit all the data available. The latter is a very crucial characteristic when relatively small samples of economic data must be tested.

However, these methods are also sensitive to noise, which can seriously affect their ability and usefulness as diagnostic tools for chaotic structure.

7.2 DESCRIPTION OF THE METHODS

7.2.1 The piecewise linear approximation method

The piecewise linear approximation method has been introduced as a prediction tool by Farmer and Sidorowitch (1987) and Casdagli (1989), but has also been developed to a tool that can help distinguishing between chaotic and stochastic alternatives [Casdagli (1991), Casdagli and Weigent (1993)]. It involves linear interpolation to construct local polynomial maps, which makes it a parametric forecasting method. According to the method, an algorithm is used[2] to fit models of the form of equation (8.2). The objective is to find the linear function $f : R^m \to R$ which gives the best prediction for x_{t+T} in a least square sense. The algorithm uses the number k of nearest neighbours as a variable smoothing parameter which, at one extreme, defines a non-linear deterministic model while at the other extreme a linear stochastic model. Small k values correspond to fitting a deterministic model while large ones to an autoregressive linear model. Intermediate k values correspond to non-linear stochastic models. Chaotic series are expected to give more accurate short-term forecasts at the deterministic extreme.

The steps of the method are the following:

1. The series is divided into two separate sets: A fitting set or a library $\{x_1,\ldots,x_{Nf}\}$, to be used to estimate the coefficients of the model, and a testing set $\{x_{Nf+1},\ldots, x_{Nf+Nt}\}$, to be used for the model evaluation.
2. A prediction time T is chosen that defines the T-step-ahead forecasting task.
3. An embedding dimension m and a delay time τ are chosen in order to embed the series in a state space (state-space reconstruction) and to construct the delayed vectors.
4. A "predictee" or a test delay vector $\mathbf{x}_t$ is chosen, with $t \geq N_f$ that will be used to impose a metric on the state space in order to find the nearest neighbours to this test delay vector.
5. The distances d_{ti} of the test vector $\mathbf{x}_t$ from the delay vectors $\mathbf{x}_i$ of the fitting set are computed for $1+(m-1)\tau \leq i \leq N_f - T$. The distances are calculated by the use of the maximum norm and the fitting set is used as a rolling library than as a fixed set, that is,

---

2 Casdagli and Weigent (1993) call this algorithm DVS (Deterministic Versus Stochastic)

it is updated each time a new point (test vector) is tested (see step 9). According to Casdagli and Weigent (1993), this can be advantageous for nonstationary systems.

6. The distances are ordered and the first k nearest neighbours $\mathbf{x}_{i(1)},\ldots,\mathbf{x}_{i(k)}$ of $\mathbf{x}_t$ are found.

7. An autoregressive model of order m is fitted to the nearest neighbours, as:

$$\mathbf{x}_{i(j)+T} \approx c_0 + \sum_{n=1}^{m} c_n \mathbf{x}_{i(j)-(n-1)\tau} \qquad , \ j = 1,\ldots,k \qquad (7.3)$$

Since this is fitted to the k nearest neighbours, the model has k equations and k takes values in the interval $2(m+1) < k < N_f - T - (m-1)\tau$.

Actually (7.3) is a linear system which in a matrix form can be written as :

$$\mathbf{A} = \mathbf{BC} \qquad (7.4),$$

where:

$$\mathbf{A} = \begin{vmatrix} x_{i(1)+T} \\ \vdots \\ x_{i(k)+T} \end{vmatrix} \quad \mathbf{B} = \begin{vmatrix} 1 & x_{i(1)+T-1} & \cdots & x_{i(1)+T-m\tau} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{i(k)+T-1} & \cdots & x_{i(k)+T-m\tau} \end{vmatrix} \quad \mathbf{C} = \begin{vmatrix} c_0 \\ \vdots \\ c_m \end{vmatrix} \qquad (7.5)$$

In (7.5) i denotes specific times in the library where the dynamics are similar to the test point. The objective is to solve for $\mathbf{C} = \mathbf{AB}^{-1}$ and estimate the coefficients $c_0,\ldots,c_m$. However, **B** is not invertible and an efficient way to solve for **C** is by factorisation of the matrix **B**. In our case this is done by an LU-decomposition [Press et. al.(1988)], although SVD analysis can also be used.

8. Once the coefficients have been estimated, the model (7.3) is used for a T-step-ahead forecast $\hat{x}_{t+T}(k)$, and an error measure is estimated. We compute the squared error, but the absolute error can also be used:

$$e_{t+T}(k) = (\hat{x}_{t+T} - x_{t+T})^2 \qquad (7.6)$$

9. A new test delay vector is chosen and steps 4 through 8 are repeated until the (t + T) runs span the whole testing set. Finally, a mean error measure is computed from the forecasts, which in our case is the Normalised Root Mean Squared Error (NRMS) defined as:

$$E_m(k) = \langle \sum_t e_{t+T}(k) \rangle^{(1/2)} / \sigma, \qquad (7.7)$$

where σ is the standard deviation of the testing set.

10. For a certain forecasting time T, the plots of the $E_m(k)$ curves are constructed as a function of both the nearest neighbours k and the embedding dimension m. The inspection of these plots provides very useful information for distinguishing between low dimensional and stochastic processes.

The forecasting ability of the model(s) has been also assessed by computing the correlation coefficient $\rho_m(k)$ between the actual and the forecasted values in the testing set.

In the piecewise linear approximation, there is a number of varying parameters that can be used to create different models, the forecasts of which can help to investigate the underlying dynamics of the series. These parameters are: 1) the smoothing parameter k which defines the nature of model (7.3) as deterministic or stochastic, 2) the embedding dimension m that defines the number of the variables and the coefficients to be estimated and 3) the forecasting time T.

The main problem of this method (which is also a problem of all the non-linear forecasting methods) is its sensitivity to noise. According to Casdagli (1991), moderate noise levels (up to 20%) can seriously affect forecastability so that no safe conclusions can be drawn about the low-dimensionality of the series, while higher noise levels (100%) can conceal even the existence of non-linearity.

Another problem is related to some properties of the series. Casdagli and Weigent (1993) show that if the data is short or non-stationary or exhibit confinement (i.e. the length of the testing set does not cover the phase portrait of the series well), forecasting results might depend on the length of the testing set and a single test might not be representative enough.

A third problem is the lack of objectively defined cut-off points of the smoothing parameter k in order to characterise the nature of the underlying dynamics as low dimensional deterministic, non-linear stochastic or linear stochastic.

7.2.2 The Simplex method

This is a non-parametric method which uses no prior information about the model used to generate the series and can be considered as a simpler variant of the more sophisticated piecewise approach, described above. However, Sugihara and May (1990) who introduced the method, claim that it is quite successful in investigating the dynamics of a time series even with a small number of observations and may also give information about its correct embedding dimension.

The steps of the method are the following:

1-6) Steps 1- 6 are identical to those of the piecewise method described above.

7) The k nearest neighbours (each one consisting of an m-dimensional delay vector) are used to create a simplex containing the predictee (the test vector) and exponential (or simple) weights $\lambda_1,...,\lambda_k$ are assigned to the distances $D_1,....,D_k$ from the test vector given by:

$$\lambda_i = \frac{e^{-D_i}}{\sum_{i=1}^{k} e^{-D_i}}, \qquad i = 1,....,k \qquad (7.8)$$

The inverse distance in (7.8) is used to assign higher weights to the closest neighbours.

8) The predicted value is obtained by projecting the domain of the simplex into its range. This means that we are keeping track of where the k nearest neighbours will be after T-time steps[3] and use the weights to determine the new position of the test vector. Numerically this is done by the simplex optimisation method [Fletcher and Reeves (1964)]. A graphical representation of the method is exhibited in Figure 7.1 below, where it is shown how the new positions of the initial nearest vectors ($\mathbf{x}_1,..,\mathbf{x}_4$) to the test vector $\mathbf{x}_p$ form a new simplex ($\mathbf{x}'_1,..,\mathbf{x}'_4$), the interior point of which corresponds to the prediction point, i.e. to the new position $\mathbf{x}'_p$ of the test vector. The exact position of the predicted point in the new simplex is determined by the use of the weights λ_i

9) The correlation coefficient $\rho_m(k)$, between the actual and the forecasted values in the testing set is computed.

---

[3] Geometrically, we track the motion of the system on its trajectory. Each nearest neighbour corresponds to a specific point on this trajectory

10) The plots of the correlation coefficient against the embedding dimension m and the prediction time T are constructed.

The simplex method is more robust than the piecewise method when small numbers of the nearest neighbours are used, i.e. more accurate predictions are obtained with low-dimensional simplex. According to Casdagli (1991) this difference is due to the fact that the piecewise method, being more sophisticated, captures more high dimensional

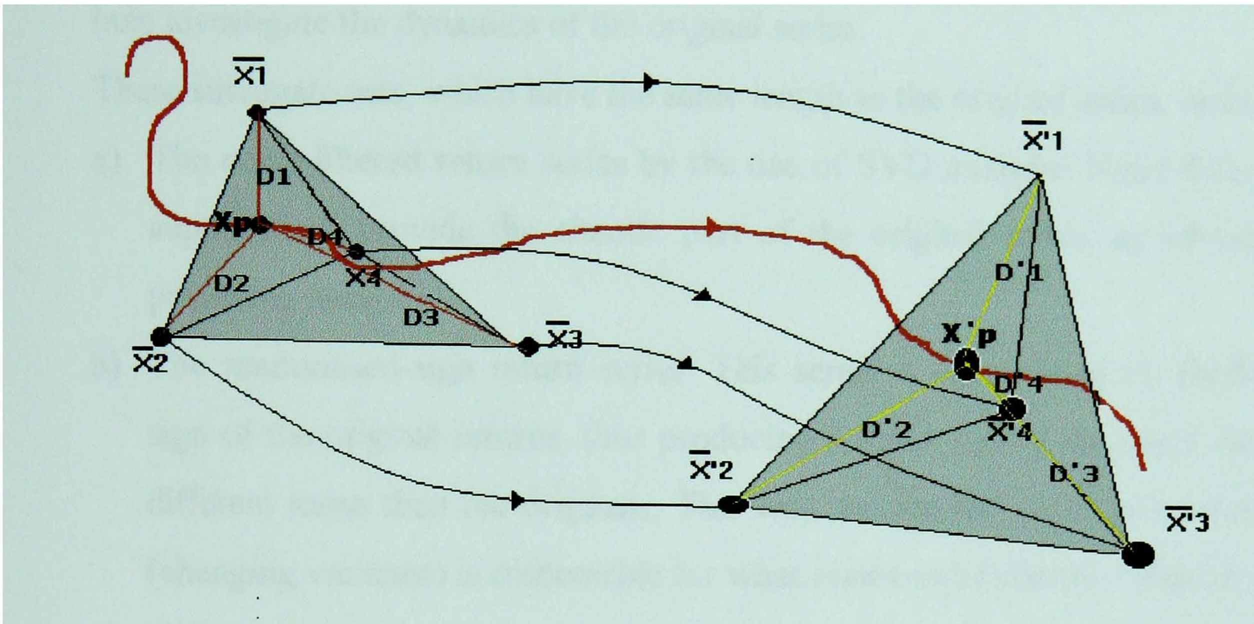

Figure 7.1. Graphical representation of the simplex method. For embedding dimension m=3, the minimum simplex is tetrahedron.

characteristics of the data and is likely to break faster as the k parameter decreases since it fits to the noise easier than the simplex method.

The practical implication is that with simplex, k is not used as a decisive parameter to the identification of the dynamics of the series. Instead it can either be a fixed number [Linden et. al. (1992), Timmermann and Satchell, (1992)] or it can be a function of the embedding dimension k = m+1, as suggested by Sughihara and May (1990). It is the latter we adopt here.

7.3 EMPIRICAL EVIDENCE

7.3.1 The DVS plot

We have used the piecewise linear approximation method as a forecasting tool and as a chaos identification tool for stock return data from the Athens' stock market. For comparison purposes we also provide forecasts for a number of surrogate sets that will help investigate the dynamics of the original series.

These surrogate sets, which have the same length as the original series, include:

a) The noise-filtered return series by the use of SVD analysis. Noise-filtered series is supposed to provide the chaotic part of the original series, as advocated in the preceding analysis.

b) The randomised-sign return series. This series is constructed by randomising the sign of the original returns, thus producing a series having the same variance but a different mean than the originals. This way we can test whether heteroscedasticity (changing variance) is responsible for what seems to be chaotic behaviour.

c) A Gaussian random series having the same mean and variance as the original series.

d) A pure chaotic series produced by a simple chaotic map, mentioned before, the logistic equation.

The basic tool to analyse the dynamics of a series using the piecewise method is to construct plots of the NRMS error curves [$E_m(k)$] as a function of both the nearest neighbours k and the embedding dimension m, for a certain prediction time T. Casdagli & Weigend (1993) call this group of curves the DVS (Deterministic versus Stochastic) plot.

We constructed these plots by calculating the average NRMS error of 100 out of sample one-step-ahead forecasts (prediction time T=1) for all the data sets mentioned above. Given the total length of the series (3181 observations), we used a fitting set of 3080 observations and a testing set of 100. This has been done for a range of nearest neighbours (k = 20,….,1500) and embedding dimensions[4] (m = 2,…,10). In order to

[4] For each k, 9 forecast runs are executed for each m = 2,..,10.

account for small data length, non-stationarity, or confinement phenomena, we repeated the calculations described above for larger lengths of the testing set (300 and 500 observations respectively). No significant changes were found with respect to the qualitative characteristics of the DVS plots, so we do not report them here.

The respective plots are shown in Figures 7.2 to 7.6. Figure 7.2 shows the plot of the logistic map, a noise free, purely chaotic specification. This is evident from the slope of the $E_m(k)$ curves which correspond to the different embedding dimensions employed. The sharply upward slope towards the linear stochastic extreme of the plot indicates that the minimum error occurs at the deterministic extreme. The minimum error curves correspond to m=3 or 4, indicating the proper embedding for the state space reconstruction of the series. Moreover, the deterministic models yield an impressive 100% forecast improvement over the linear stochastic ones, as measured on the minimum error curves. Notice that forecast improvement is deteriorating with increasing m.

It is much more difficult to interpret Figure 7.3, which shows the original return series plot. The $E_m(k)$ curves exhibit a slightly upward slope towards the stochastic extreme and the lowest error curve for small k values corresponds to m=6, which has been found to be an upper limit to the dimensionality of the underlying attractor of the series. This is an indication of a possible deterministic explanation, yet, the lowest error for most of the curves corresponds to intermediate k values, indicating a high dimensional deterministic or a non-linear stochastic explanation.

The picture becomes more obscure since forecasting improvement over the linear stochastic models is marginal (5%-8%), measured on the lower error curves, and even the best of the forecasts in terms of the NRMS error are extremely poor and always greater than 1. However, the noise inherent to these series could be responsible for poor forecastability.

The latter is verified by the DVS plot of the noise-filtered series in Figure 7.4. Forecasts have dramatically improved (more than 60%) compared to the original series' ones. The $E_m(k)$ curves are moving upward towards the right-end of the plot, indicating a 30% forecast improvement of the left-end deterministic and nonlinear stochastic models

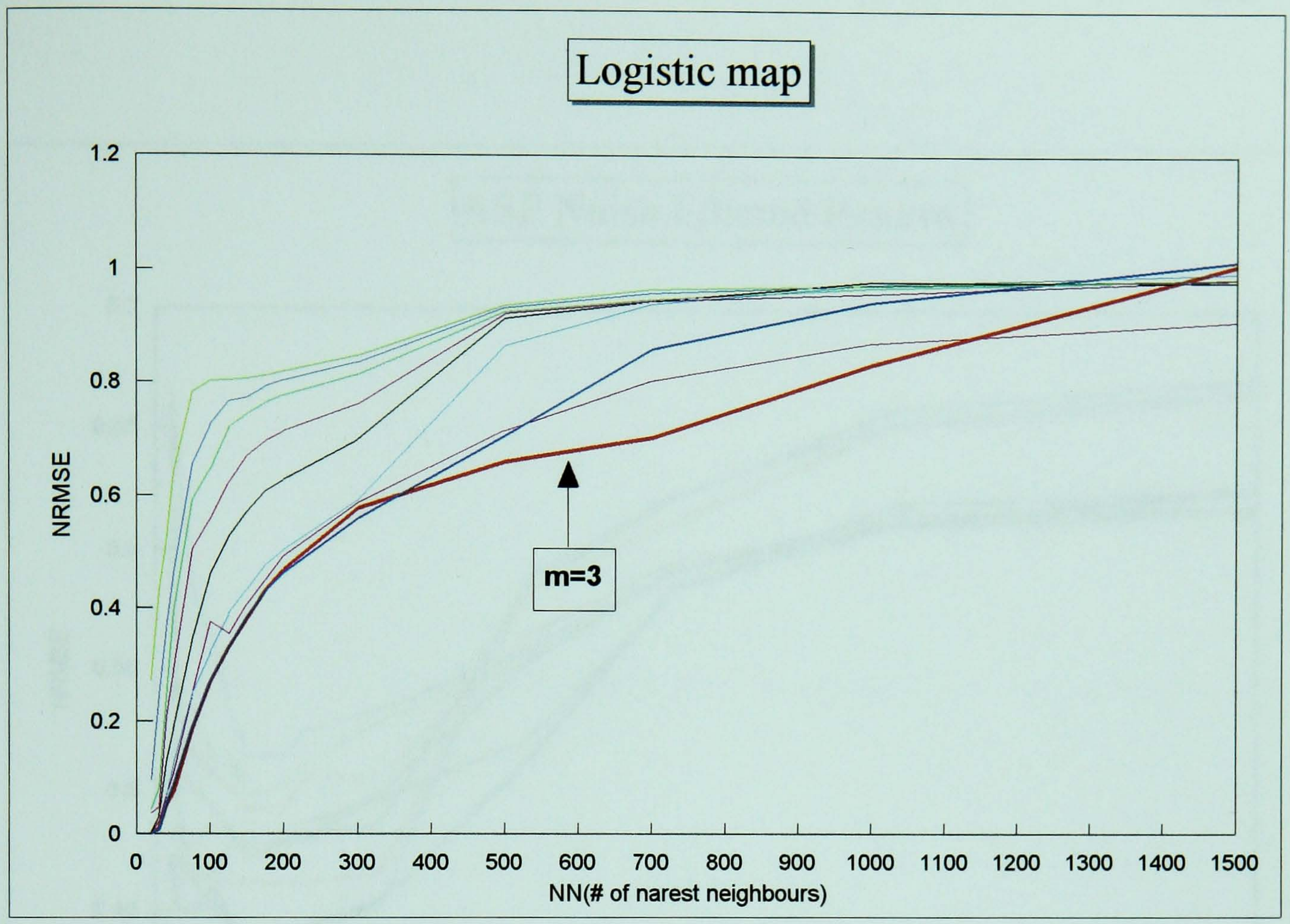

Figure 7.2DVS plot of the logistic map for m=2,...10. Minimum error occurs at the extreme left, indicating the chaotic nature of the logistic map

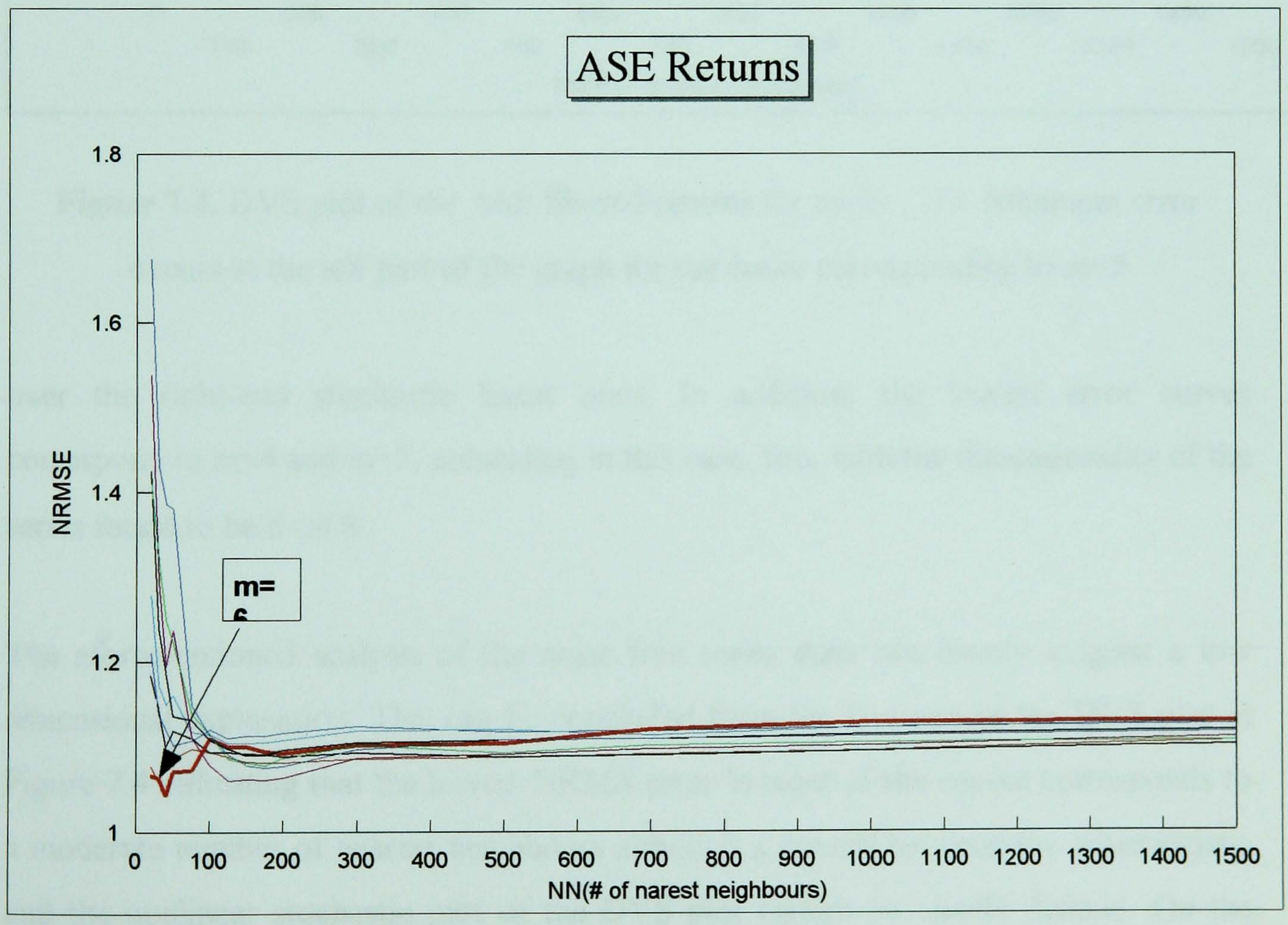

Figure 7.3DVS plot of the ASE returns for m=2,...10. Minimum error occurs at the extreme left for the curve corresponding to m=6.

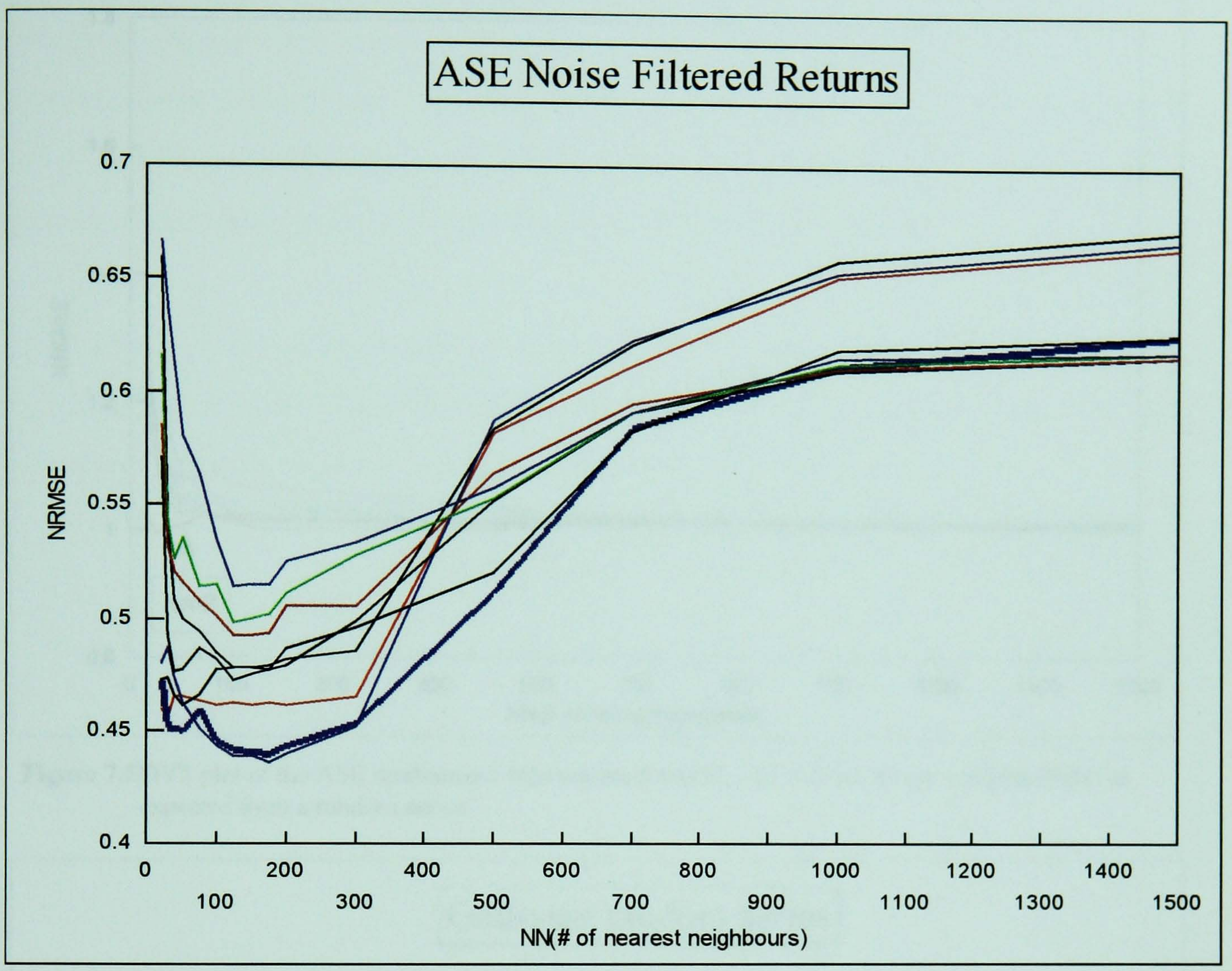

Figure 7.4. DVS plot of the ASE filtered returns for m=2,...,10. Minimum error occurs at the left part of the graph for the curve corresponding to m=5

over the right-end stochastic linear ones. In addition, the lowest error curves correspond to m=4 and m=5, coinciding in this case, too, with the dimensionality of the series found to be d ≅4.8.

The aforementioned analysis of the noise free series does not clearly suggest a low dimensional explanation. This can be concluded from the U-shape of the DVS plot in Figure 7.4 indicating that the lowest NRMS error in most of the curves corresponds to a moderate number of nearest neighbours although a cut-off between the deterministic and the nonlinear stochastic part of the DVS plot cannot be clearly defined. On the other hand, the minimum NRMS error for the lowest error curve (m=5) occurs at the extreme left of the plot, an indication of a chaotic explanation.

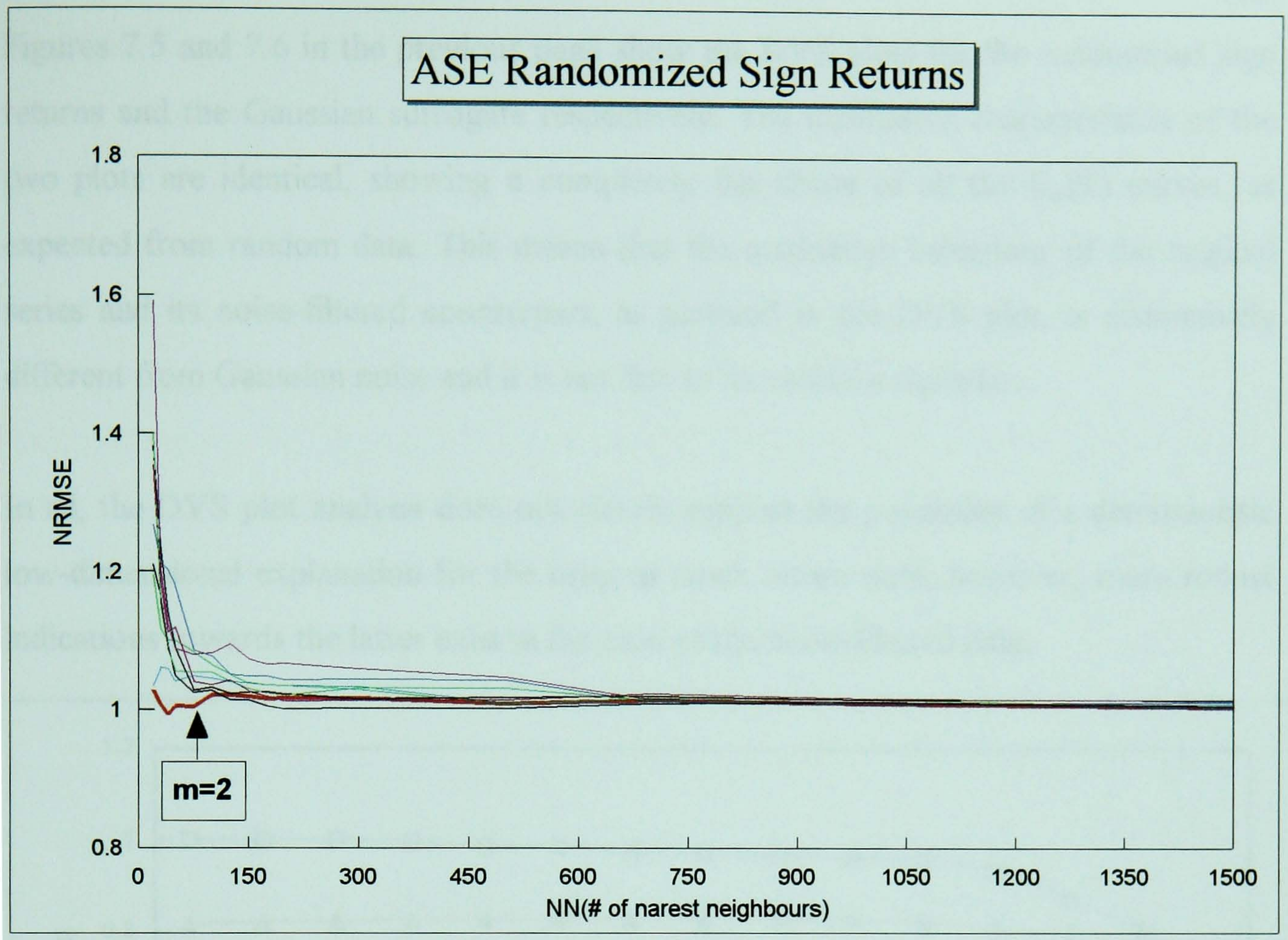

Figure 7.5DVS plot of the ASE randomized sign returns for m=2,...10. All curves are completely flat as expected from a random series.

Figure 7.6DVS plot of the Gaussian random surrogate of the ASE returns for m=2,...10. All curves are flat and erratic, showing no apparent structure.

Figures 7.5 and 7.6 in the previous page show the DVS plots for the randomised sign returns and the Gaussian surrogate respectively. The qualitative characteristics of the two plots are identical, showing a completely flat shape of all the $E_m(k)$ curves, as expected from random data. This means that the qualitative behaviour of the original series and its noise-filtered counterpart, as pictured in the DVS plot, is distinctively different from Gaussian noise and it is not due to its variance signature.

In all, the DVS plot analysis does not clearly support the possibility of a deterministic low-dimensional explanation for the original stock return data; however, more robust indications towards the latter exist in the case of the noise-filtered data.

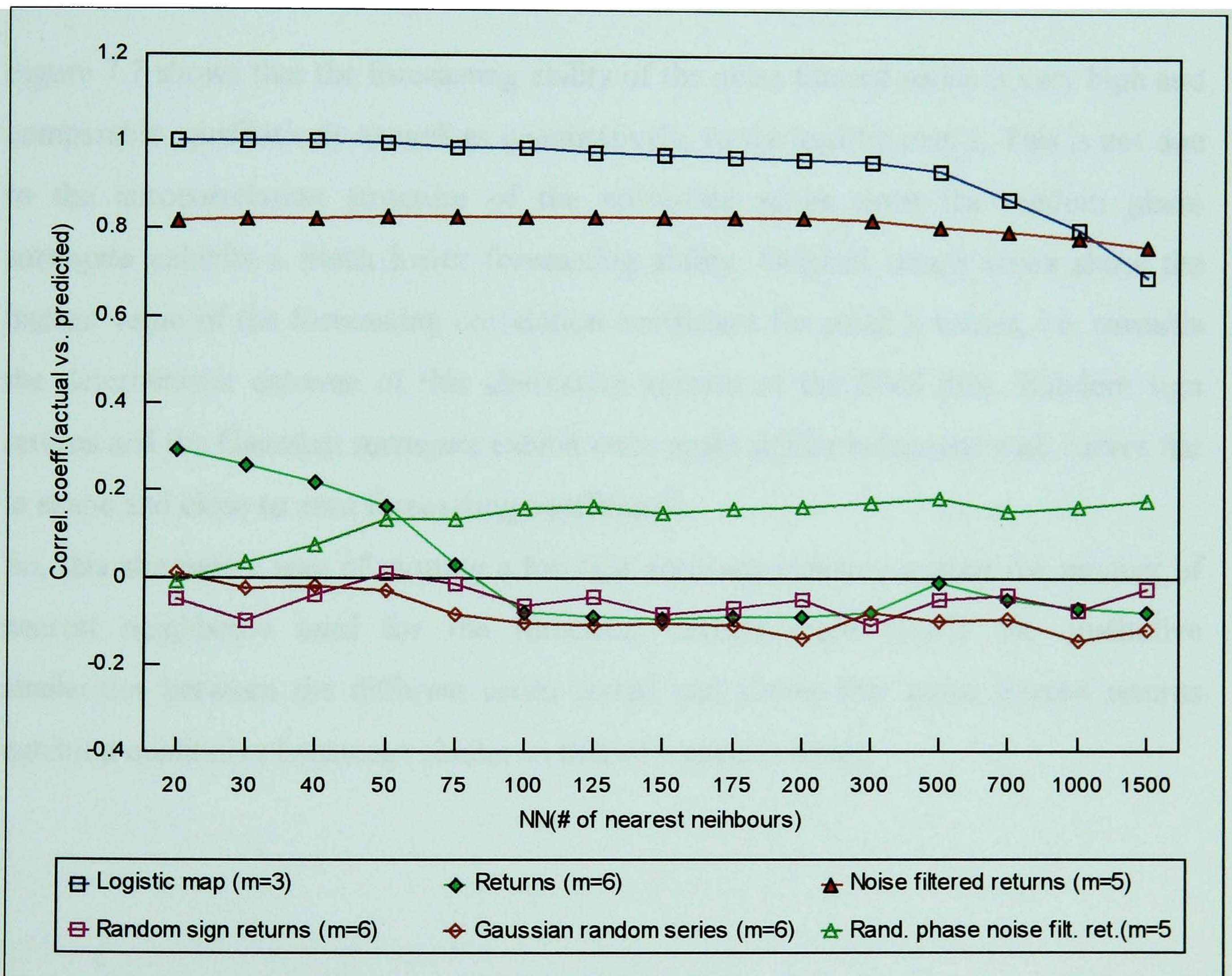

Figure 7.7. Alternative version of the DVS plot, using correlation coefficient between actual and predicted values instead of the NRMS error.

This is shown in Figure 7.7 above where instead of the NRMS error the correlation coefficient between actual and predicted values is calculated[5] and is plotted against the number of nearest neighbours.

The m value we used to generate the forecasts in this plot is the best embedding value, i.e. the one which produces the lower forecast error, as indicated by the DVS analysis for the logistic map, the original return series and the noise filtered series. The m values employed for each series are specified in the legend box of Figure 7.7.

This is done for the original series and all the surrogates, including an additional one, i.e. the randomised-phase noise filtered returns, having the same autocorrelation as the original noise filtered series. This way we can test whether the qualitative behaviour and forecastability of the noise-filtered series is due to its autocorrelation structure.

Figure 7.7 shows that the forecasting ability of the noise filtered series is very high and comparable, qualitatively as well as quantitatively, to the logistic map's. This is not due to the autocorrelation structure of the noise-free series since the random phase surrogate exhibits a much lower forecasting ability. Original return series show the highest value of the forecasting correlation coefficient for small k values, i.e. towards the deterministic extreme of this alternative version of the DVS plot. Random sign returns and the Gaussian surrogate exhibit once again similar behaviour with curves flat in shape and close to zero forecasting coefficients.

So, this alternative way of plotting a forecast accuracy measure against the number of nearest neighbours used for the forecasts, reveals more clearly the qualitative similarities between the different series tested and shows that noise filtered returns exhibit a qualitative behaviour similar to that of a chaotic series.

---

5 Correlation coefficient is defined as :

$$r(X_i,F_i) = \frac{1}{n-1}\sum_{i=1}^{n}\frac{\left(X_i-\overline{X}\right)}{S_X}\frac{\left(F_i-\overline{F}\right)}{S_F}$$

where $i = 1,\ldots,n$ is the length of the testing sample, X_i is the actual datum for time period i and F_i is the forecast for the same period and S_X and S_F correspond to the standard deviation of the actual and predicted values respectively The r value ranges from -1 to +1 for perfect negative and positive autocorrelation respectively.

7.3.2 The "varying prediction time" and the "dimensionality" techniques

Another way of detecting low deterministic dynamics and distinguishing them from noise is suggested by Sugihara and May (1990). According to them the simplex method, presented above, can be used to generate forecasts with varying prediction time T. Then the plot of the correlation coefficient between actual and predicted values (r) versus prediction time T can be used to detect chaotic behaviour. Chaotic series are expected to show a decreasing r with increasing T [6], while a random noise process should not exhibit such dependence. This is what we call the "varying prediction time" technique. In addition they show that the same method can be used to specify an upper bound to the dimensionality of the attractor of a putatively chaotic system.
According to the "dimensionality" technique, the correlation measure r, is sensitive to the choice of m (the embedding dimension). In the case of a chaotic system, the m value corresponding to the highest r in the r vs. m plot indicates an upper bound on the attractor's dimensionality.

The results of the "varying prediction time" technique are presented in Figure 7.8 (a-c). All the series used in the DVS approach are presented here, as well, and the m parameter used for each series is the same as the one used to generate the predictions presented in Figure 7.7.
The original return series in Figure 7.8a show indeed a decrease of r with T, but after the third prediction step the curve flattens out. The maximum r is a little higher than 0.2, indicating a very strong noisy component[7] but, despite the noise, this technique seems to catch the r vs. T dependence expected from a series with a chaotic component. The r vs. T dependence is much more sound in the case of the noise filtered series where a clear decaying r vs. T relationship is pictured as expected from a chaotic

---

[6] Such a behaviour is a characteristic feature of chaos reflecting the presence of a positive Lyapunov exponent, with the magnitude of the exponent related to the rate of decrease of r with T.

[7] Sugihara & May (1990), suggest that the maximum r value could be also considered as an indication of the additive noise inherent to a series.

"VARYING PREDICTION TIME" TECHNIQUE (I)

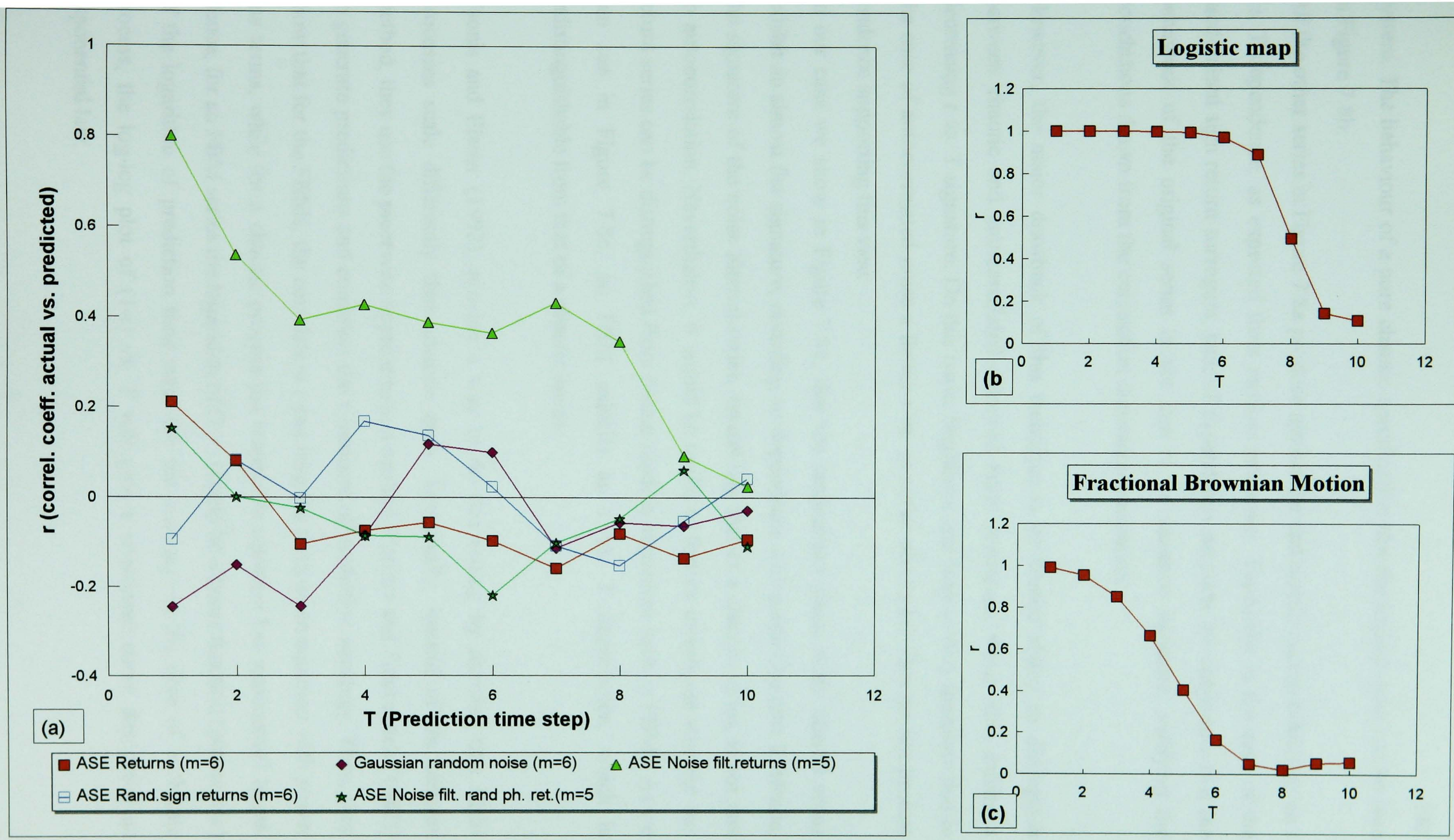

Figure 7.8 (a-c) Detection of chaotic behaviour through the investigation of the relationship between correlation coefficient and prediction time, for different series
Graph a, presents the ASE series and related surrogates, graph b, a chaotic process and graph c, an FBM process

system. The behaviour of a pure chaotic specification like the logistic map can be seen in Figure 7.8b.

All the other series in Figure 7.8a produce almost flat and erratic curves exhibiting no r vs. T dependence, as expected from random processes. Especially in the case of the randomised sign return surrogate, lack of dependence supports the conjecture that the behaviour of the original series is not due to its variance signature, verifying the conclusions drawn from the correlation dimension estimation.

However, the major drawback of this technique is its limited ability to distinguish between chaotic and autocorrelated coloured noise processes, which also exhibit a decreasing r vs. T signature. On this issue, Sughihara and May (1990) speculate that in the case of autocorrelated noise a flatter r vs. m curve occurs but they do not present evidence supporting this view.

In our case we show in Figure 7.8a, that the randomised phase noise filtered series exhibit an almost flat signature, revealing no dependence with prediction time T. Hence, the signature of the noise filtered series, resembling that of a chaotic series, is not due to autocorrelation. Nevertheless, it would be useful to further investigate whether our return series can be distinguished from fractal random sequences such as FBMs. As we can see in Figure 7.8c, an FBM exhibits an r vs. T dependence, which is indistinguishable from that of a chaotic series.

Tsonis and Elsner (1992), provide a way to do this testing by showing that FBM processes scale differently than chaotic ones. Specifically, instead of the simplex method, they use the piecewise approximation method [Farmer and Sidorowich (1987)] to generate predictions and construct the r measure of prediction accuracy. Then, they show that for the FBMs, the correlation r can take the form of an equation with power law terms, while for a chaotic process the respective equation has exponential terms. Hence, for an FBM series the logarithm of (1-r) should be a linear function (power law) of the logarithm of prediction time step. On the contrary, in the case of a chaotic process, the log-log plot of (1-r) vs. T will give a non-linear curve describing an exponential law.

To perform this testing we also used the piecewise linear method[8], and the series tested include, as in the previous case, the original and noise filtered returns, all the surrogate series used previously, the logistic map and an FBM series with H = 0.65.

Figures 7.9b and 7.9c present the log-log plots of (1-r) vs. T for the Logistic map and the FBM series. It can be clearly seen that the behaviour of both series is the expected one from a chaotic or a coloured noise sequence, respectively.

In Figure 7.9a, the behaviour of the return series and its surrogates are presented. This time, the original return series shows no different signature than the surrogate series. With respect to the latter, no one exhibits the behaviour expected either from a chaotic or from an FBM specification. In fact their behaviour is indistinguishable from the one of the random Gaussian surrogate. However, the noise-filtered returns are quite different exhibiting once more the behaviour expected from a chaotic series.

In conclusion, the "varying prediction time" approach supports a chaotic explanation only for the noise filtered return series while it gives indications of chaotic behaviour for the original return series.

---

8 The simplex method has been employed too to perform this test and see whether the results differ. In fact, the results were identical and this is the reason we do not report them here.

"VARYING PREDICTION TIME" TECHNIQUE (II)

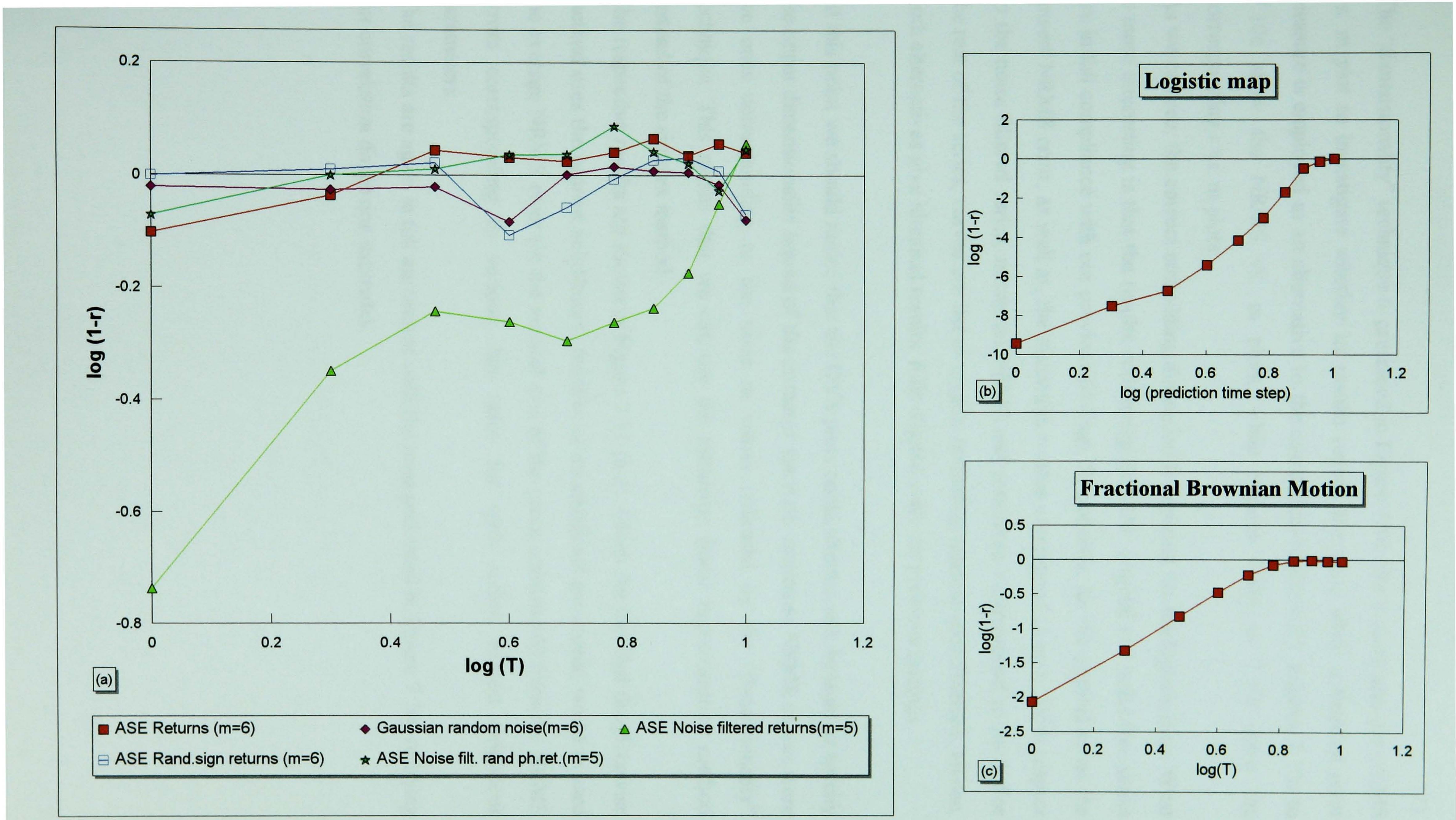

Figure 7.9 (a-c) Log-log plot of (1-r) vs. T for different series. This test can distinguish between chaotic specifications and FBMs. In graph a, the ASE reries and the related surrogates are presented. In graph b, a chaotic process and in graph c, an FBM process.

The “dimensionality” technique is presented in Figure 7.10. We include also the NRMS vs. m plot to investigate whether its results remain the same when a forecast error measure is employed as an alternative to the correlation measure (r). Figures 7.10a to 7.10e show the NRMS vs. m plots, while Figures 7.10f to 7.10j show the corresponding r vs. m plots.

As we can see, the correct embedding dimension is depicted for the logistic map. What is more interesting is that the results concerning the ASE original and noise free series are in full compliance with our previous findings. Specifically, for the original series the lowest NRMS error, as well as, the maximum r-value correspond to m=6. With respect to the noise filtered series, minimum NRMS and maximum r correspond to m=5. For the rest of the series, curves are flat or erratic, indicating random specifications. Hence, both alternatives give identical results, fully aligned with our previous findings.

At this point we should notice that the DVS plots could alternatively be used to specify the upper dimensionality bound of the attractor since the minimum NRMS curves were the ones corresponding to the same m values indicated by the “dimensionality” technique. This means that we can use the piecewise linear approximation method instead of the simplex method.

The respective results are shown in Figure 7.11 (a-e). Given the fact that the piecewise method uses the nearest neighbours’ number as an additional parameter, we have used the average NRMS error (in the vertical axis of the plot), constructed from the NRMS errors corresponding to various, but same for each series, nearest neighbour parameters.

The results are again in full agreement with the ones presented in Figure 7.10, verifying our correlation dimension estimates.

"Dimensionality" technique (Simplex method)

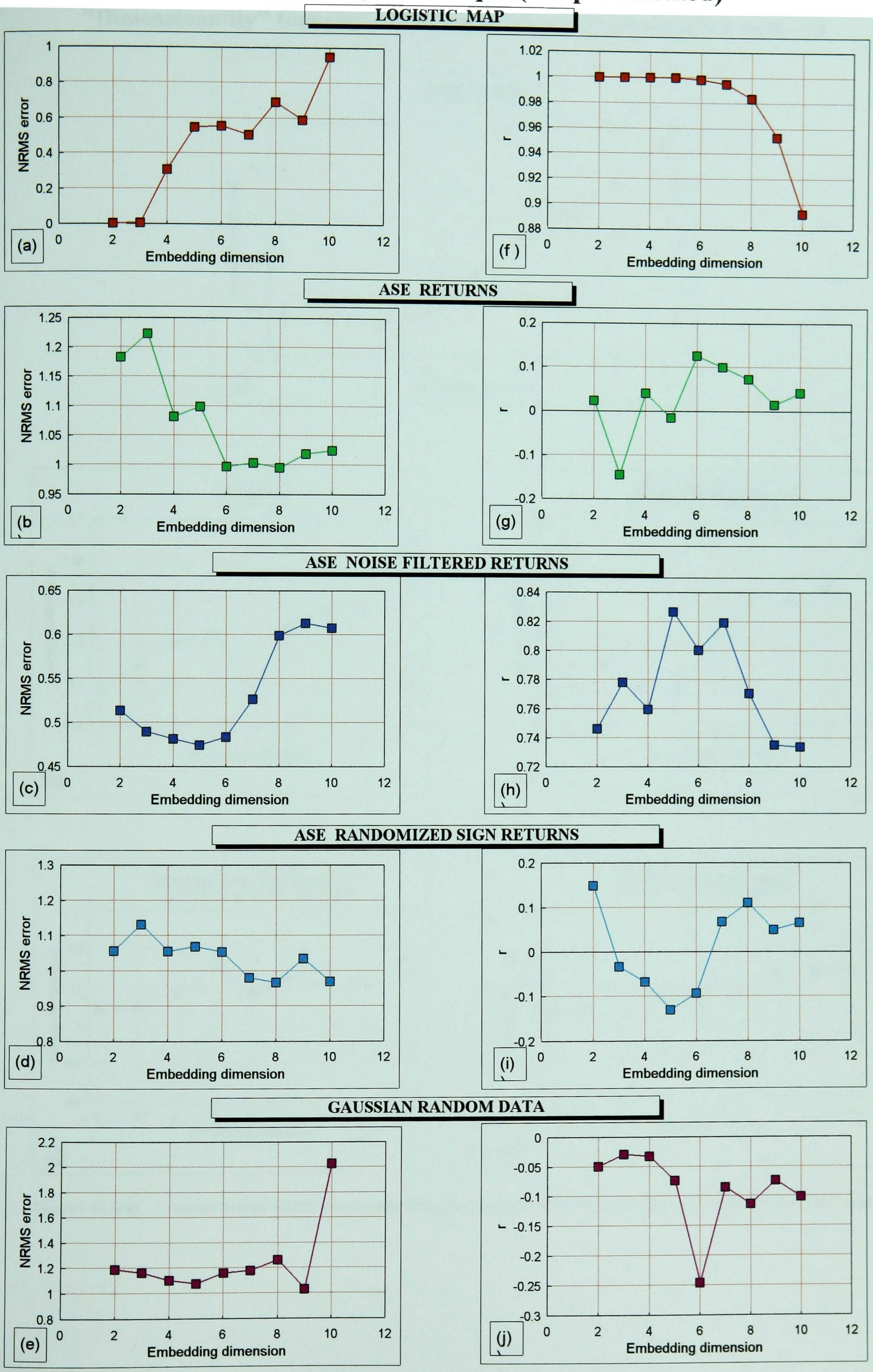

Figure 7.10 (a-e) Forecast NRMSE vs. embedding dimension for differents series, as a test for an upper dimensionality bound
Figure 7.10 (f-j) Correlation coefficient between actual & predicted values vs. embedding dimension for differents series as a test for an upper dimensionality bound.

"Dimensionality" technique (Piecewise linear approximation method)

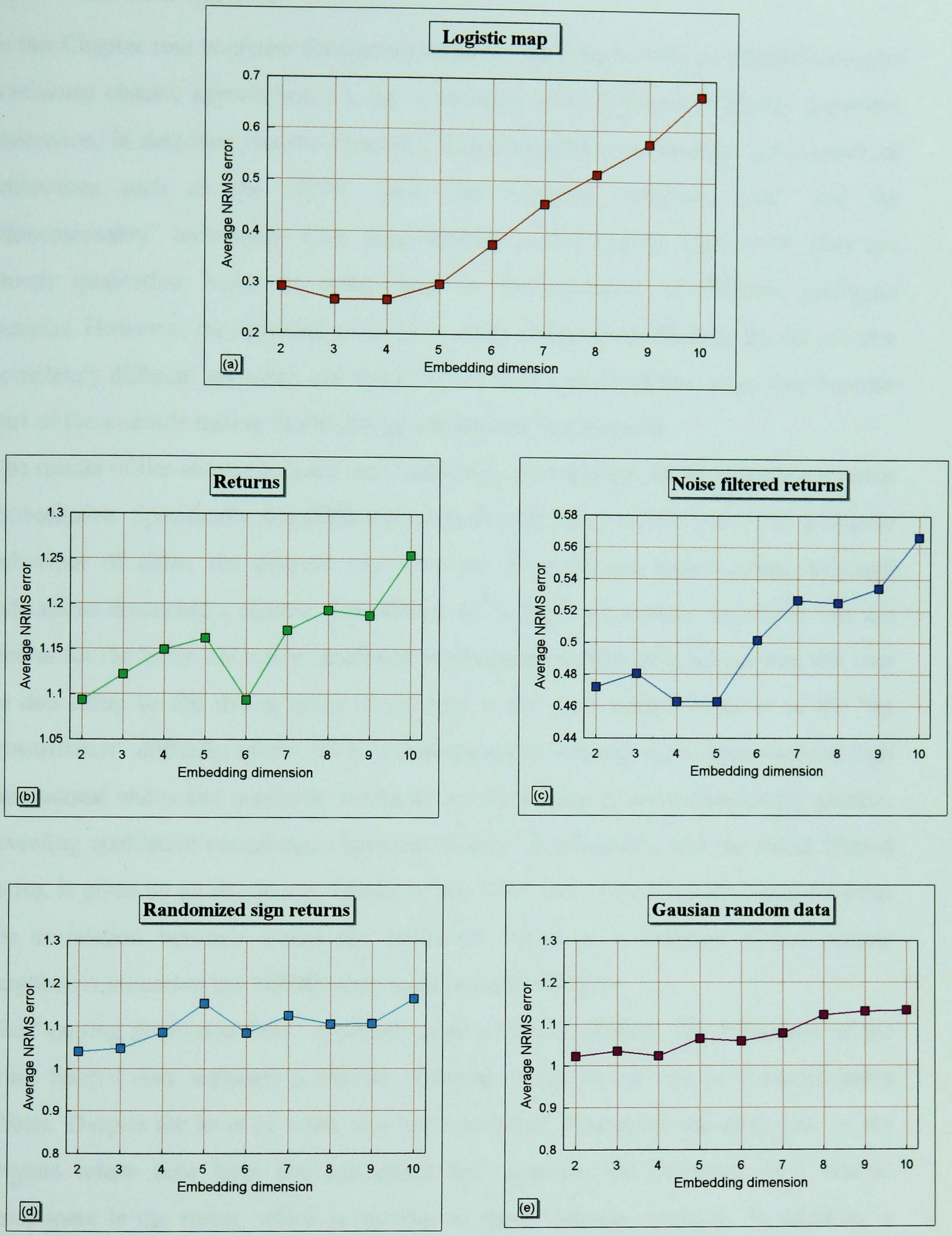

Figure 7.11 (a-e) Average forecast NRMS error vs. embedding dimension for different series as a test for an upper dimensionality bound

7.4 CONCLUSIONS

In this Chapter two nonlinear forecasting methods are employed as an alternative to the traditional chaotic approaches, i.e. the correlation dimension and Lyapunov exponent estimation, in detecting chaotic dynamics. These methods can be used in a number of techniques such as the "DVS" plot, the "varying prediction time" and the "dimensionality" techniques. Like most methods in the chaotic framework, they are mostly qualitative, based on comparisons in the behaviour of different surrogate samples. However, they provide a means to verify the previous findings by the use of a completely different approach and this is where their usefulness lies, since they become part of the multiple testing methodology adopted by this research.

The results of the above tools are not conflicting; nevertheless, some of them are rather inconclusive. Specifically, the DVS approach does not give a clear picture of a chaotic behaviour of either the original return or the noise filtered return series, although indications favouring a chaotic explanation, as opposed to a linear stochastic one are clearer for the latter. However, nonlinear stochasticity cannot be ruled out and this may be due either to the strong noise component in the ASE return series or to the "by construction" difficulty of the DVS in distinguishing between noisy chaos or/and high dimensional chaos and nonlinear stochastic specifications. A more illuminating picture, revealing qualitative resemblance between chaotic specifications and our noise filtered series, is given by an alternative version of the DVS plot. This alternative plot presents the correlation between actual and predicted values as a function of the nearest neighbours instead of the NRMS error used in the DVS plots.

The "varying prediction time" approach gives a clearer picture. The behaviour of the noise filtered data supports a chaotic explanation that is not due to autocorrelation effects. Despite the level of noise, this test can better distinguish the behaviour of the original return data from the surrogates and supports the existence of a chaotic component in the series, which is not due to their variance signature. In addition, a useful version of this technique was able to exclude the possibility of FBM behaviour for the noise-filtered returns.

Finally, the "dimensionality" approach, performed by using both the nonlinear prediction methods, verified our previous findings regarding the correlation dimension estimation and SVD analysis for both the original and the noise filtered ASE returns.

Chapter 8

SHORT TERM FORECASTING AND ECONOMIC ASSESSMENT OF FORECASTS

8.1 INTRODUCTION

In the previous Chapter we employed non-linear forecasting techniques to further test our series for the existence of chaotic components and to verify the results of the multiple testing framework adopted in this study.

In this Chapter we do not analyse the structure and the generating mechanism of our series. Instead, we focus on short-term predictability issues and assess the forecasting ability of alternative linear and non-linear techniques by the use of various measures of forecasting accuracy.

Moreover, a very interesting question is addressed, i.e. whether these alternative techniques are useful in terms of generating real economic results or in other words, whether they are applicable in a profitable trading strategy. So, a new, quite useful criterion for comparing various forecasting models is adopted.

In the light of the evidence of the preceding chaotic analysis and the differences and similarities found between the ASE and the LSE Stock Index return data, it is interesting to see whether this is also reflected in this exercise, as well.

Specifically, since a chaotic explanation could not be ruled out for the ASE series and chaotic series can be short-term predictable, it is interesting to see whether the nonlinearities found in our data are exploitable by non-liner methods in a short-term prediction procedure.

On the other hand, no indication of a chaotic structure was found for the LSE data, which, however, seems to be non-linear. Since both the algorithms presented in the previous Chapter allow for exploiting stochastic components in a time series, as well, we apply them to the LSE series to assess their prediction power in comparison to that of the ASE series.

In the recent literature there is mixed evidence on whether nonlinearities found in economic series are exploitable for forecasting purposes. Several economic series have been tested for nonlinear forecast improvements, such as exchange rates by Diebold and Nason (1990), Meese and Rose (1989) and Mizrach (1989), stock returns by LeBaron (1991) and White (1988) and gold series by Prescott and Stengos (1988). In all of this literature nonparametric estimation techniques were employed and no out-of-sample forecast improvement was found. In most of the cases stated above, nonlinear forecasts were found to be no better than naïve random walk forecasts in terms of the standard forecast error measures such as Mean Absolute Error (MAE) or Mean Squared Error (MSE). One exception to these findings can be found in LeBaron (1992a,b), who uses S&P stock index and exchange rate data to show that forecast improvements are possible for certain "predictability pockets" characterised by low volatility.

However, significance of nonlinearity does not imply its economic importance and vice-versa. Satchell and Timmermann (1992a) have shown that forecast error measures such as MAE and MSE might not be adequate criteria to assess the economic significance of nonlinear forecasts when the former is based on sign predictions. This is so, because the probability of predicting the sign of a stochastic variable need not be a decreasing function of MAE or MSE, if the predicted value and prediction error are not independent, a particularly relevant situation in nonlinear predictions.

Hence, it is possible to construct nonlinear predictions with large MAE and MSE but capable in producing substantial profits when implemented in a trading strategy, which exploits only the sign prediction in a series.

This approach can be found in Timmermann and Satchell (1992) and Satchell and Timmermann (1992b), who assess the economic significance of nonlinear forecasts through a trading strategy for various stock indices and find superiority of the nonlinear forecasts against a linear autoregressive alternative.

8.2 DESCRIPTION OF METHODOLOGY

We use the two stock index daily return series that have been used as the basic data sets to be analysed in this study. The first one is from the ASE market, spanning the period from January 1981 to October 1993 with 3181 observations (the one used for the analysis in the preceding Chapters). The second one is from the LSE market, covering exactly the same time period with 3347 observations, in order to have directly comparable results between the two series.

Our assessment of the forecasting accuracy is based on 4 different measures, namely the Mean Absolute Error, the Normalised Root Mean Squared Error, the Theil's U statistic and the correlation coefficient between actual and predicted values.

For a testing set of length $i = 1,\ldots,n$ where X_i is the actual datum for time period i and F_i is the forecast for the same period, the error is defined as:

$$e_i = X_i - F_i \tag{8.1}$$

The Mean Absolute Error (MAE) is defined as:

$$MAE = \frac{\sum_{i=1}^{n} |e_i|}{n} \tag{8.2}$$

This is a standard statistical measure, assigning equal weight to all errors.

The Normalised Root Mean Squared Error (NRMSE) is defined as:

$$NRMSE = \frac{\sqrt{\frac{\sum_{i=1}^{n} e_i^2}{n}}}{\sigma} \tag{8.3}$$

where σ is the standard deviation of the testing set.

This is a version of the Mean Squared Error which considers the disproportionate cost of the large errors, yet, normalisation by the use of std., creates a relative basis for comparison between different data sets. Specifically, NRMSE = 1, indicates that our forecast is no better than predicting the mean of the testing set. Accordingly, NRMSE > 1 indicates forecast worse than predicting the mean and NRMSE < 1 indicates a better, than predicting the mean, forecast.

Theil's U statistic is another relative measure allowing for the comparison of the forecasting method employed with a naïve forecast. It also gives larger weight to large errors and is defined as:

$$U = \sqrt{\frac{\sum_{i=1}^{n-1}\left(\frac{F_{i+1} - X_{i+1}}{X_i}\right)^2}{\sum_{i=1}^{n-1}\left(\frac{X_{i+1} - X_i}{X_i}\right)^2}} \qquad (8.4)$$

Notice that the bracketed term in the numerator of (8.4) corresponds to the forecasted relative change, while the respective term in the denominator corresponds to the actual relative change. A value of $U = 1$, indicates that the forecasting technique employed is as good as the naïve method. A value of $U < 1$, indicates that the forecasting technique used is better than the naïve method and a value of $U > 1$, indicates the opposite.

The correlation coefficient between actual and predicted values has been defined in the previous Chapter, and, as it is well known, takes values in the range [-1,+1].

Our assessment of the economic significance of short-term predictions is based on a trading strategy known as ***"switching rule"***. This switching strategy was originally described by Alexander (1961) and Fama and Blume (1966) and has been more recently adopted by Timmermann and Satchell (1992), Pesaran and Timmermann (1992) and Satchell and Timmermann (1992a,b). According to this rule, the investor holds the index portfolio when the index return is predicted to rise, otherwise he holds cash. This active trading strategy is compared to a "buy and hold" passive strategy, where the investor buys the index portfolio and holds it during the whole prediction period.

The switching portfolio has no bankruptcy risk, since the investor is not allowed to go short and its payoffs can be directly compared to those of the market portfolio in a mean-variance sense, since no gearing is involved in the investment strategy.

It should be noticed that holding cash in the case of the switching strategy does not give the optimal payoff of the switching portfolio, since the money could be invested overnight. Yet, the "holding cash" hypothesis serves better the purpose of assessing the possible excess profit return as the outcome of the market timing skills due to the adopted forecasting model[1].

A second hypothesis maintained by our method is that the market index is tradable on a daily basis. Stock indices, unlike Futures, are not directly tradable. So, this hypothesis might not be always realistic in real market conditions. However, our purpose of comparing the methods in economic terms, is not seriously affected.

The switching strategy has been simulated with the use of a set of recursive equations conditioned upon the predicted sign, which describe the changes in the portfolio value over time. Sign predictions form an indicator variable:

$$
\begin{aligned}
I_{t+1} &= 1 \text{ if } {}_tR_{t+1} \geq 0 \text{ or} \\
&\quad 0 \text{ if } {}_tR_{t+1} < 0
\end{aligned}
\qquad (8.5)
$$

where ${}_tR_{t+1}$ is the daily return forecast at time t, for the t+1 period, generated by the two non-linear nearest neighbour models presented in the previous Chapter.

In our methodology, the investor's initial portfolio wealth is 100 \$. The portfolio wealth is calculated for both strategies on a daily basis throughout the prediction period, at the end of which we get the end portfolio wealth for each strategy.

In addition, for comparison purposes we use a switching strategy based on:

- a naïve random walk prediction, according to which the best return forecast for period t is the actual return at t-1.
- a moving average forecast based on the 20 last values [MA(20)], found to be very popular among the analysts of the Greek market.
- a linear autoregressive model, the lag of which is determined for each data set by partial autocorrelograms and Schwartz's information criterion [Schwartz (1978)], whichever gives the longest lag. Like the nonlinear models, this model uses a rolling library of adjustable length.

---

[1] It could be argued that our "switching" trading strategy does not take into account the strength of the forecasts as portfolio theory would suggest. Actually, our trading model is able to handle "filter rules" that can exploit the forecast value and not its sign only. Alternatively, the equity participation in the case of a positive forecast sign could be adjusted according to how strong this forecast is, in order to reduce equity exposure in the case of a weak positive sign. However, our experiments with different versions of both these alternatives gave no consistently better results than our basic approach, so we do not report them here.

Hence, for each data set, we calculate and compare the End Wealth from 6 different portfolios, the market portfolio and 5 switching portfolios, corresponding to the two strategies:

- The Buy & Hold strategy (market portfolio)
- A switching strategy based on forecasts generated by the piecewise approximation method
- A switching strategy based on forecasts generated by the simplex method
- A switching strategy based on forecasts generated by a random walk prediction
- A switching strategy based on forecasts generated by an MA(20) model
- A switching strategy based on forecasts generated by an AR model

As we shall show in the Tables of the next section, the prediction sets (or portfolios) are much more than 6 for each of the two data series, due to the parameters involved in the nonlinear prediction models. Specifically, for both models, we generate predictions for various embedding dimensions (m) ranging from 2 to 10. In addition, as presented in the previous Chapter, the piecewise method involves a varying parameter for the number of nearest neighbours (NN) used in the prediction process. This time four different parameter values have been selected, namely 20, 50, 100 and 200, covering a wide range of nearest neighbours.

For each set of predictions a rolling library of varying length has been used[2] to test the sensitivity of predictions against different prediction periods and different library lengths. To this end, five scenarios are constructed for each data set, depending on a library length of 1 to 5 years respectively. For each scenario the prediction period covers the remaining of the data set[3].

For the LSE case in particular, much more past data is available, so it was feasible to assess the impact of using a much longer library on the prediction results. As we shall

---

[2] It is obvious that the rolling library is used only in the case of the two nonlinear and the linear AR model for the fitting process. In the case of the naïve forecast and the MA model, there is no need for a rolling library.

[3] Each scenario exploits the total observations in each data set. For example, both data sets span a period of approximately 13 years. The first scenario uses a rolling library of 1 year and a 12-year prediction period, the second a rolling library of 2 years data and an 11-year prediction period, etc.

present in the next section in detail, we fix a prediction period of 11 and 8 years respectively, and use the longest library available to generate predictions.

The second part of this exercise is a sub-period analysis. We introduce two different criteria, volatility and sign change frequency and we try to investigate whether any of these criteria is related to predictability. That is, whether high or low volatility and/or sign change frequency can help in identifying "predictability pockets". An analytical presentation of the methodology and the issues addressed in this second part of our analysis will follow in the relevant section.

The empirical evidence of the first part will be discussed in the next section, and one Table is presented for each prediction period, incorporating the results by all the different forecasting models.

For each prediction model, a complete set of indicators is presented, as:

- **Forecasting Accuracy Statistical Measures**

1. Normalised Root Mean Squared Error (NRMSE)
2. Mean Absolute Error (MAE)
3. Correlation Coefficient (CC) between actual and predicted values and a t-value to assess its statistical significance.
4. Theil's U test

- **Indicators related to the "switching strategy"**

5. The **Mean** of the switching portfolio
6. The **Standard Deviation** of the switching portfolio

   Both measures above are used to compare the alternative strategies and prediction models in a "mean-variance" sense[4].
7. The **correct sign prediction** percentage
8. The **correlation coefficient** between actual and predicted sign and a t-value to assess its statistical significance. Sign prediction is the criterion adopted to determine

---

[4] Alternatively we could use the "Sharpe ratio" $[(R_e\text{-}R_f)/\sigma]$ to compare the performance of the different models against the buy and hold strategy, but it was very difficult to have estimates of R_f on a daily basis for the ASE market due to inadequacies of the statistical data infrastructure. However, as we shall see in the empirical part of this application, our conclusions are not qualitatively affected since in most cases the one strategy is better than the other in terms of both the mean and the std.

the asset allocation in the switching strategy. A percentage well above 50% indicates a successful trading strategy and is directly related to the End Wealth value.

9. The **End Wealth** value of the switching portfolio, discussed above, assuming no transaction costs. This could be particularly relevant when trading in a futures index or in FX markets where transaction costs are fairly small. However, in stock markets, transaction costs are important and should be considered in order to correctly assess the economic value of our forecasts. To this end, an additional indicator is introduced, the break-even transaction cost, presented below.
10. The **excess wealth,** produced by the switching strategy over that of the buy & hold strategy assuming no transaction costs.
11. The **switching frequency,** which measures the percentage change in the asset allocation, i.e. how many times during the forecasting period the prediction model gives a sign to switch between the "hold index" and the "hold cash" positions. Actually, the switching frequency corresponds to the realised transactions and determines the total transaction cost.
12. The **break-even transaction cost**, which gives the same mean return on the market and the switching portfolios or the cost at which the End Wealth values of the market and the switching portfolios become equal.

    Transaction costs are particularly important for switching portfolios based on daily signals due to the high turnover on such portfolios. In the past [Alexander (1961), Fama and Blume (1966)], transaction costs have been proven to make filter rules economically non-exploitable and market efficiency invincible.

    By comparing the real transaction cost in each market to the break even transaction cost, the possibility of beating the market by using forecasts is assessed. This can be related to the market efficiency definition by Jensen (1968), according to whom, in an efficient market with respect to information set Ω_t, it is impossible to make economic profits by trading on the basis of this information set. However, our aim is to assess and compare the economic value of different forecasts than to test market efficiency. The latter is a much more difficult task as Fama (1991) has pointed out, since our results will be joined tests of the model generating expected returns and the Efficient Market Hypothesis.

Transaction costs might give an intuitive explanation of serial correlation in the stock markets. As Satchell and Timmermann (1992a) indicate, serial correlation in returns may not be removed by arbitrage due to high transaction costs, especially when the stock index is broadly based, containing many small and illiquid shares.

The effective transaction cost for the Greek market is estimated to be on average 0.50% - 0.55% for the period under study. Notice that from 1995 this cost has dropped to 0.25% - 0.35%. The latter is due to the removal of a State-controlled mechanism imposing the transaction cost level to the market.

The transaction cost of the U.K market according to Beckers (1992) varies from 0.15% (the net commission rate) to 0.65%, if taxes will be considered.

Indicators related to the "buy and hold Strategy"

For the Market portfolio (i.e. the "benchmark" portfolio in our analysis) reflecting the **"buy and hold Strategy"**, which corresponds to each prediction period, the Mean, Standard Deviation and End Wealth are estimated.

8.3 EMPIRICAL EVIDENCE

8.3.1 ASE & LSE prediction results (Library:1-5 years, Prediction period: 12-8 years)

Tables 8.1 to 8.5 present the five **ASE data** scenarios corresponding to rolling libraries of 1 to 5 years and prediction periods of 12 to 8 years respectively.

Each Table[5] is divided into 4 major areas: one in the upper left part for the linear prediction results, one in the upper right part for the market index results, one in the middle part for the results of the nonlinear predictions (piecewise model) and one in the lower part for the simplex model. Notation wise, hereafter, the alternative models will be referred to in their abbreviated form as RW (Naïve Random Walk model), MA (Moving Average model), AR (Linear autoregressive model), PW (Piecewise approximation model), SX (Simplex model).

Due to the density of information in each Table, the most important indicators, namely the best value of each forecast accuracy measure among all different prediction portfolios, the highest correct sign proportion ratio, the best end wealth value, and the break- even transaction cost exceeding the average estimated transaction cost for each market, appear in bold and are shaded. In addition, the best values for the same indicators for each one of the two nonlinear models appear in bold to make comparisons easier. Statistical measures in Tables 8.1 to 8.5 show indications of predictability in the ASE returns. This is more evident in terms of the correlation coefficient between realised and predicted values, which is found to be significant in the case of the three linear specifications and in the best prediction cases of the nonlinear models. Comparing the prediction models in terms of statistical measures, we can see in the same Tables that the linear models outperform the nonlinear ones, irrespectively of the length of the library and the prediction period employed. Specifically, the linear AR(2) model selected in the ASE case, exhibits the best performance in terms of the NRMSE and MAE measures in all the different prediction periods, but its superiority is marginal (improvement does not exceed 3%) in comparison to the best error estimates obtained by the two nonlinear models.

---

[5] All prediction results hereafter will be presented in the same way described above to make comparisons easier.

Table 8.1 : ASE Prediction Results (Library : 1 year, Prediction period : 12 years)

Linear Predictions

Prediction model	Random Walk	MA(20)	Linear model
NRMSE	1.213431	1.011392	**0.977839**
MAE	0.011475	0.010296	**0.009677**
Correl. Coeff. Act.vs.pred.	**0.2638**	0.1079	0.2576
t-value	14.28	5.84	13.94
Theil's U	1.000	0.897	0.904
Mean	0.002458	0.001573	0.002347
Std. Dev.	0.013917	0.012683	0.014643
Correct Sign Proportion	**64.07%**	57.69%	63.02%
Correl. Coef. sign prediction	0.2805	0.1518	0.2580
t-value	15.19	8.22	13.97
End wealth predict. (EW)	**132,811**	9,990	95,815
EW(pred) vs. EW(b&h)	8484.99%	545.78%	6093.53%
Switching frequency	0.35892	0.05902	0.35005
Break even trans. cost	0.42%	**1.07%**	0.40%

Market Index (Buy & hold strategy)	
Mean	0.0009355
Std. Dev.	0.0178616
End Wealth	1,547

Non Linear predictions (Piecewise approx. method)

nn = 20 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.092349	1.122882	1.138767	1.251678	1.376019	1.466239	1.476407	1.576273	1.679267
MAE	0.010537	0.010882	0.010995	0.011890	0.012455	0.012870	0.013579	0.014425	0.015168
Correl. Coeff. Act.vs.pred.	0.0480	0.0599	0.0467	0.0664	0.0389	0.0710	0.1294	0.0634	0.0153
t-value	2.60	3.24	2.53	3.59	2.11	3.85	7.01	3.43	0.83
Theil's U	1.030	1.023	1.228	0.936	0.948	0.879	1.218	1.289	1.853
Mean	0.001487	0.001242	0.001344	0.001361	0.001105	0.001205	0.001451	0.001327	0.001120
Std. Dev.	0.016102	0.014151	0.014167	0.015478	0.014149	0.014106	0.013625	0.013851	0.014181
Correct Sign Proportion	57.90%	57.80%	57.73%	57.08%	56.33%	56.60%	59.37%	56.57%	55.99%
Correl. Coef. sign prediction	0.1600	0.1606	0.1578	0.1423	0.1298	0.1369	0.1926	0.1394	0.1274
t-value	8.66	8.69	8.54	7.70	7.03	7.41	10.43	7.55	6.90
End wealth predict. (EW)	7,773	3,795	5,109	5,374	2,536	3,406	6,982	4,868	2,657
EW(pred) vs. EW(b&h)	402.45%	145.34%	230.23%	247.38%	63.95%	120.19%	351.31%	214.70%	71.72%
Switching frequency	0.33504	0.37871	0.39236	0.39406	0.40259	0.41419	0.40873	0.43876	0.44012
Break even trans. cost	0.17%	0.08%	0.10%	0.11%	0.04%	0.07%	0.13%	0.09%	0.04%

nn = 50 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.037962	1.055514	1.046847	1.086963	1.067221	1.173846	1.220511	1.267309	1.260715
MAE	**0.010208**	0.010376	0.010355	0.011088	0.010560	0.012047	0.012402	0.013431	0.013006
Correl. Coeff. Act.vs.pred.	0.0873	0.0949	**0.0985**	0.0080	0.0905	-0.0748	0.0214	-0.0627	-0.0006
t-value	4.73	5.14	5.33	0.43	4.90	-4.05	1.16	-3.40	-0.03
Theil's U	0.943	0.989	1.059	1.025	1.044	1.061	0.913	0.999	1.233
Mean	0.001861	0.001816	0.001659	0.000838	0.001666	0.000496	0.000902	0.000164	0.000709
Std. Dev.	0.015458	0.014815	0.015405	0.014221	0.014496	0.013406	0.013726	0.012663	0.013471
Correct Sign Proportion	**61.51%**	60.97%	60.29%	53.33%	59.84%	51.28%	52.71%	48.86%	50.87%
Correl. Coef. sign prediction	0.2286	0.2195	0.2075	0.0673	0.1979	0.0282	0.0548	-0.0169	0.0198
t-value	12.38	11.88	11.23	3.65	10.71	1.53	2.97	-0.91	1.07
End wealth predict. (EW)	**23,211**	20,293	12,825	1,164	13,104	427	1,403	162	798
EW(pred) vs. EW(b&h)	1400.36%	1211.75%	729.04%	-24.78%	747.05%	-72.41%	-9.28%	-89.56%	-48.42%
Switching frequency	0.31457	0.36643	0.33572	0.50802	0.40293	0.51314	0.55544	0.49130	0.54384
Break even trans. cost	0.29%	0.24%	0.22%	0.00%	0.18%	0.00%	0.00%	0.00%	0.00%

nn = 100 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.027828	1.024828	1.035555	1.038290	1.059817	1.053578	1.064913	1.070327	1.072033
MAE	0.010565	0.010567	0.010624	0.010676	0.010856	0.010821	0.010948	0.010998	0.011025
Correl. Coeff. Act.vs.pred.	-0.0775	-0.0712	-0.0793	-0.0741	-0.1052	-0.0780	-0.0948	-0.1005	-0.1014
t-value	-4.20	-3.86	-4.29	-4.01	-5.70	-4.22	-5.13	-5.44	-5.49
Theil's U	0.965	0.973	0.879	0.902	0.930	0.831	0.749	**0.705**	0.714
Mean	0.000273	0.000189	0.000451	0.000429	0.000279	0.000458	0.000179	0.000295	0.000075
Std. Dev.	0.014161	0.014411	0.013454	0.014580	0.013626	0.013377	0.013145	0.013590	0.013529
Correct Sign Proportion	48.31%	47.63%	49.98%	49.27%	49.30%	49.71%	48.65%	48.82%	48.00%
Correl. Coef. sign prediction	-0.0295	-0.0429	0.0005	-0.0138	-0.0112	-0.0046	-0.0247	-0.0217	-0.0385
t-value	-1.59	-2.32	0.03	-0.75	-0.61	-0.25	-1.34	-1.17	-2.08
End wealth predict. (EW)	223	174	374	351	226	382	169	237	125
EW(pred) vs. EW(b&h)	-85.61%	-88.76%	-75.81%	-77.30%	-85.39%	-75.30%	-89.08%	-84.67%	-91.95%
Switching frequency	0.43262	0.47697	0.39918	0.40737	0.44217	0.47902	0.45513	0.43398	0.43262
Break even trans. cost	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

nn = 200 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.022978	**1.019098**	1.032778	1.035219	1.044977	1.040305	1.052818	1.043154	1.043559
MAE	0.010531	0.010517	0.010585	0.010632	0.010690	0.010658	0.010821	0.010717	0.010712
Correl. Coeff. Act.vs.pred.	-0.0578	-0.0398	-0.0659	-0.0680	-0.0808	-0.0477	-0.0847	-0.0420	-0.0399
t-value	-3.13	-2.15	-3.57	-3.68	-4.37	-2.58	-4.59	-2.27	-2.16
Theil's U	1.009	1.016	0.925	0.934	0.936	0.859	0.787	0.718	0.741
Mean	0.000251	0.000349	0.000546	0.000536	0.000487	0.000709	0.000281	0.000548	0.000522
Std. Dev.	0.013897	0.014313	0.013320	0.014517	0.014194	0.014230	0.013994	0.014360	0.014177
Correct Sign Proportion	48.11%	48.41%	50.19%	49.68%	49.98%	50.77%	49.30%	50.02%	49.64%
Correl. Coef. sign prediction	-0.0288	-0.0222	0.0096	-0.0012	0.0053	0.0203	-0.0096	0.0042	-0.0040
t-value	-1.56	-1.20	0.52	-0.07	0.29	1.10	-0.52	0.23	-0.22
End wealth predict. (EW)	209	278	494	480	417	797	227	498	461
EW(pred) vs. EW(b&h)	-86.51%	-82.05%	-68.04%	-68.95%	-73.07%	-48.46%	-85.30%	-67.80%	-70.19%
Switching frequency	0.35960	0.37052	0.34459	0.36370	0.39372	0.39099	0.42170	0.41965	0.40805
Break even trans. cost	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Non Linear predictions (Simplex method)

m =	2	3	4	5	6	7	8	9	10
NRMSE	1.033684	1.013836	1.006229	**0.994938**	1.008780	1.001974	1.002229	1.002553	1.020103
MAE	0.010675	0.010418	0.010405	**0.010237**	0.010422	0.010264	0.010273	0.010445	0.010803
Correl. Coeff. Act.vs.pred.	0.0648	0.0846	0.0854	**0.1214**	0.0412	0.0521	0.0520	0.0554	0.0097
t-value	3.51	4.58	4.62	6.57	2.23	2.82	2.82	3.00	0.53
Theil's U	1.002	1.232	1.161	1.030	1.109	0.989	1.109	**0.909**	1.477
Mean	0.001088	0.001441	0.001259	0.001457	0.000746	0.001102	0.000936	0.000906	0.000003
Std. Dev.	0.014778	0.014595	0.014111	0.015070	0.012711	0.014127	0.013043	0.015997	0.007687
Correct Sign Proportion	**56.81%**	56.74%	55.48%	56.40%	53.12%	55.71%	55.78%	52.92%	48.38%
Correl. Coef. sign prediction	0.1364	0.1351	0.1164	0.1317	0.0764	0.1085	0.1200	0.0461	-0.0077
t-value	7.39	7.32	6.30	7.13	4.14	5.87	6.50	2.49	-0.42
End wealth predict. (EW)	2,416	6,785	3,985	**7,103**	889	2,517	1,550	1,421	101
EW(pred) vs. EW(b&h)	56.18%	338.56%	157.56%	359.12%	-42.55%	62.71%	0.16%	-8.17%	-93.48%
Switching frequency	0.38008	0.40873	0.32344	0.36370	0.32276	0.36677	0.40089	0.20505	0.17093
Break even trans. cost	0.04%	0.12%	0.10%	0.14%	0.00%	0.05%	0.00%	0.00%	0.00%

Table 8.2 : ASE Prediction Results (Library : 2- year, Prediction period : 11- year)

Linear Predictions

Prediction model	Random Walk	MA(20)	Linear model
NRMSE	1.215589	1.011381	**0.972928**
MAE	0.012152	0.010851	**0.010163**
Correl. Coeff. Act.vs.pred.	0.2612	0.1075	**0.2630**
t-value	13.52	5.57	13.62
Theil's U	1.000	0.889	0.731
Mean	0.002608	0.001681	0.002419
Std. Dev.	0.014458	0.013144	0.015028
Correct Sign Proportion	**64.42%**	58.04%	63.04%
Correl. Coef. sign prediction	0.2867	0.1573	0.2568
t-value	14.85	8.15	13.29
End wealth predict. (EW)	**107,329**	9,010	64,708
EW(pred) vs. EW(b&h)	6814.94%	480.51%	4068.99%
Switching frequency	0.35509	0.06043	0.36927
Break even trans. cost	0.44%	**1.08%**	0.38%

Market Index (Buy & hold strategy)	
Mean	0.0011229
Std. Dev.	0.0186782
End Wealth	1,552

Non Linear predictions (Piecewise approx. method)

nn = 20 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.302388	1.271631	1.318286	1.432281	1.627142	1.693308	1.764105	1.873055	2.057911
MAE	0.011642	0.012264	0.012878	0.013807	0.014842	0.015723	0.016599	0.017287	0.018839
Correl. Coeff. Act.vs.pred.	0.0378	0.0289	0.0763	0.0581	0.0179	0.0575	0.0737	0.0391	0.0147
t-value	1.96	1.49	3.95	3.01	0.93	2.98	3.82	2.02	0.76
Theil's U	1.072	1.478	0.815	0.781	0.957	1.693	0.837	0.808	0.861
Mean	0.001696	0.001263	0.001476	0.001388	0.001215	0.001380	0.001466	0.001537	0.001310
Std. Dev.	0.015513	0.014853	0.013985	0.016011	0.014765	0.015737	0.014528	0.014316	0.014407
Correct Sign Proportion	59.98%	57.78%	58.49%	57.74%	56.28%	57.22%	58.37%	57.93%	56.51%
Correl. Coef. sign prediction	0.1967	0.1539	0.1697	0.1526	0.1248	0.1432	0.1676	0.1592	0.1299
t-value	10.19	7.97	8.79	7.90	6.46	7.42	8.68	8.24	6.73
End wealth predict. (EW)	9,365	2,941	5,200	4,107	2,585	4,026	5,069	6,119	3,338
EW(pred) vs. EW(b&h)	503.36%	89.45%	235.05%	164.62%	66.57%	159.36%	226.58%	294.20%	115.05%
Switching frequency	0.35882	0.39239	0.39090	0.42521	0.43044	0.44610	0.42372	0.44013	0.45132
Break even trans. cost	0.19%	0.06%	0.12%	0.09%	0.04%	0.08%	0.10%	0.12%	0.06%

nn = 50 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.007135	1.044430	1.091801	1.104824	1.083890	1.147429	1.207521	1.241956	1.266153
MAE	0.010492	0.010874	0.011074	0.011543	0.011101	0.012150	0.012902	0.013760	0.013521
Correl. Coeff. Act.vs.pred.	**0.1612**	0.1267	0.0793	0.0605	0.1795	0.0704	0.1336	0.0689	0.1302
t-value	8.35	6.56	4.11	3.13	9.29	3.65	6.92	3.57	6.74
Theil's U	0.976	1.050	0.856	0.939	0.759	1.080	1.033	**0.494**	0.872
Mean	0.002057	0.001843	0.001528	0.001277	0.001995	0.001104	0.001202	0.001176	0.001403
Std. Dev.	0.016393	0.015107	0.014247	0.014135	0.014895	0.015054	0.014096	0.013146	0.014605
Correct Sign Proportion	**62.63%**	60.31%	59.83%	56.58%	60.87%	54.90%	54.94%	52.74%	54.42%
Correl. Coef. sign prediction	0.2492	0.2045	0.1961	0.1355	0.2180	0.0988	0.0989	0.0663	0.0910
t-value	12.90	10.59	10.15	7.02	11.29	5.12	5.12	3.43	4.71
End wealth predict. (EW)	**24,621**	13,870	5,973	3,054	20,859	1,921	2,499	2,330	4,274
EW(pred) vs. EW(b&h)	1486.25%	793.60%	284.86%	96.74%	1243.89%	23.74%	61.03%	50.13%	175.34%
Switching frequency	0.31033	0.36852	0.36404	0.43640	0.37598	0.46624	0.55949	0.48937	0.54457
Break even trans. cost	**0.33%**	0.22%	0.14%	0.06%	0.26%	0.02%	0.03%	0.03%	0.07%

nn = 100 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.004262	1.003955	1.015518	1.015088	2.558955	1.024821	1.026310	1.033138	1.039209
MAE	0.010820	0.010818	0.010826	0.010877	0.022441	0.011053	0.011108	0.011183	0.011257
Correl. Coeff. Act.vs.pred.	0.0754	0.0856	0.0453	0.0544	0.0620	0.0469	0.0519	0.0320	0.0138
t-value	3.90	4.43	2.34	2.82	3.21	2.43	2.69	1.66	0.71
Theil's U	1.010	0.965	0.939	0.948	3.061	0.794	0.726	0.689	0.743
Mean	0.000705	0.000787	0.000954	0.001308	0.001263	0.001138	0.001160	0.000978	0.000756
Std. Dev.	0.015594	0.014733	0.013665	0.014497	0.015268	0.014238	0.014729	0.014660	0.014711
Correct Sign Proportion	52.07%	51.73%	54.08%	54.49%	55.09%	53.23%	54.35%	52.56%	51.66%
Correl. Coef. sign prediction	0.0422	0.0411	0.0834	0.0932	0.0994	0.0663	0.0886	0.0520	0.0333
t-value	2.19	2.13	4.32	4.83	5.15	3.43	4.59	2.69	1.73
End wealth predict. (EW)	660	824	1,285	3,317	2,942	2,108	2,235	1,370	757
EW(pred) vs. EW(b&h)	-57.46%	-46.92%	-17.19%	113.74%	89.52%	35.79%	44.02%	-11.71%	-51.26%
Switching frequency	0.38046	0.43790	0.39164	0.40656	0.46997	0.43491	0.41701	0.42894	0.41999
Break even trans. cost	0.00%	0.00%	0.00%	0.07%	0.05%	0.03%	0.03%	0.00%	0.00%

nn = 200 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.002433	**1.001170**	1.012046	1.011683	1.016640	1.016240	1.023623	1.019732	1.023362
MAE	0.010825	0.010810	**0.010807**	0.010851	0.010927	0.010943	0.011048	0.011016	0.011072
Correl. Coeff. Act.vs.pred.	0.0844	0.0988	0.0599	0.0685	0.0594	0.0717	0.0462	0.0710	0.0554
t-value	4.37	5.11	3.10	3.55	3.08	3.71	2.39	3.68	2.87
Theil's U	1.064	1.003	0.985	1.015	0.949	0.843	0.779	0.673	0.740
Mean	0.000725	0.000839	0.001067	0.001432	0.001207	0.001259	0.001112	0.001029	0.000959
Std. Dev.	0.015245	0.014457	0.013512	0.014260	0.014311	0.014550	0.015112	0.014844	0.014604
Correct Sign Proportion	51.62%	52.22%	54.79%	54.72%	54.76%	54.23%	53.97%	52.89%	52.37%
Correl. Coef. sign prediction	0.0397	0.0583	0.1023	0.1028	0.1015	0.0908	0.0823	0.0611	0.0506
t-value	2.06	3.02	5.30	5.32	5.26	4.70	4.26	3.16	2.62
End wealth predict. (EW)	697	946	1,739	4,622	2,530	2,910	1,966	1,573	1,304
EW(pred) vs. EW(b&h)	-55.08%	-39.07%	12.05%	197.82%	63.03%	87.46%	26.68%	1.34%	-16.00%
Switching frequency	0.31705	0.34987	0.35658	0.37673	0.38642	0.38717	0.42298	0.42223	0.41179
Break even trans. cost	0.00%	0.00%	0.01%	0.11%	0.05%	0.06%	0.02%	0.00%	0.00%

Non Linear predictions (Simplex method)

m =	2	3	4	5	6	7	8	9	10
NRMSE	1.009662	1.003906	0.992231	**0.988525**	0.995102	0.988574	0.991069	0.994341	1.011523
MAE	0.010858	0.010851	0.010764	**0.010697**	0.010816	0.010667	0.010712	0.010905	0.011298
Correl. Coeff. Act.vs.pred.	0.1564	0.1415	0.1604	**0.1604**	0.1300	0.1512	0.1395	0.1369	0.0958
t-value	8.10	7.32	8.31	8.30	6.73	7.83	7.22	7.09	4.96
Theil's U	1.561	1.238	1.074	1.294	1.236	0.889	1.108	**0.593**	1.650
Mean	0.001429	0.001583	0.001408	0.001420	0.001383	0.001550	0.001487	0.001234	0.000549
Std. Dev.	0.015408	0.015878	0.013913	0.015583	0.014312	0.015957	0.013584	0.017858	0.008828
Correct Sign Proportion	**58.78%**	58.49%	55.99%	55.50%	55.17%	57.81%	57.22%	54.31%	49.98%
Correl. Coef. sign prediction	0.1710	0.1665	0.1280	0.1140	0.1189	0.1491	0.1479	0.0765	0.0475
t-value	8.85	8.62	6.63	5.90	6.16	7.72	7.66	3.96	2.46
End wealth predict. (EW)	4,587	**6,919**	4,335	4,478	4,059	6,335	5,358	2,718	435
EW(pred) vs. EW(b&h)	195.51%	345.77%	179.33%	188.48%	161.49%	308.13%	245.18%	75.13%	-72.00%
Switching frequency	0.38642	0.40731	0.35360	0.38269	0.32301	0.36106	0.38568	0.18128	0.19023
Break even trans. cost	0.10%	0.14%	0.11%	0.10%	0.11%	0.15%	0.12%	0.12%	0.00%

Table 8.3 : ASE Prediction Results (Library : 3- year, Prediction period : 10- year)

Linear Predictions

Prediction model	Random Walk	MA(20)	Linear model
NRMSE	1.218042	1.012433	**0.972659**
MAE	0.013176	0.011752	**0.010969**
Correl. Coeff. Act.vs.pred.	0.2582	0.1019	**0.2671**
t-value	12.60	4.97	13.04
Theil's U	1.000	0.889	0.805
Mean	0.002864	0.001879	0.002562
Std. Dev.	0.015258	0.013847	0.014638
Correct Sign Proportion	**64.01%**	57.24%	63.25%
Correl. Coef. sign prediction	0.2755	0.1331	0.2537
t-value	13.44	6.49	12.38
End wealth predict. (EW)	90,045	8,692	**94,744**
EW(pred) vs. EW(b&h)	3936.69%	289.68%	4147.35%
Switching frequency	0.35951	0.06552	0.38891
Break even trans. cost	**0.43%**	**0.87%**	**0.40%**

Market Index (Buy & hold strategy)	
Mean	0.001306
Std. Dev.	0.0195817
End Wealth	2,231

Non Linear predictions (Piecewise approx. method)

nn = 20 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.568005	1.438469	1.465939	1.558598	1.570425	1.573801	1.630205	1.756906	1.876516
MAE	0.013669	0.013932	0.014502	0.015557	0.016003	0.016328	0.017203	0.018251	0.019197
Correl. Coeff. Act.vs.pred.	-0.0162	-0.0078	0.0231	0.0323	0.0194	0.0402	0.0929	0.0876	0.0721
t-value	-0.79	-0.38	1.13	1.58	0.95	1.96	4.54	4.27	3.52
Theil's U	1.115	1.314	0.818	1.516	0.865	1.922	1.451	0.940	1.640
Mean	0.001688	0.001487	0.001614	0.001444	0.001236	0.001231	0.001520	0.001415	0.001428
Std. Dev.	0.016452	0.015175	0.015366	0.015536	0.015305	0.015670	0.016125	0.015931	0.015902
Correct Sign Proportion	59.22%	58.67%	58.13%	57.58%	54.72%	54.89%	55.90%	55.65%	54.72%
Correl. Coef. sign prediction	0.1782	0.1747	0.1618	0.1491	0.0915	0.0943	0.1180	0.1113	0.0912
t-value	8.69	8.53	7.90	7.28	4.47	4.60	5.76	5.43	4.45
End wealth predict. (EW)	5,533	3,426	4,635	3,096	1,888	1,865	3,709	2,892	2,982
EW(pred) vs. EW(b&h)	148.06%	53.59%	107.81%	38.77%	-15.38%	-16.41%	66.29%	29.66%	33.70%
Switching frequency	0.38639	0.40823	0.39731	0.42503	0.43175	0.44771	0.44687	0.45107	0.48047
Break even trans. cost	0.10%	0.04%	0.08%	0.03%	0.00%	0.00%	0.05%	0.02%	0.03%

nn = 50 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.183395	1.183022	1.148265	1.198496	1.200652	1.193749	1.244681	1.382998	1.293175
MAE	0.012105	0.012305	0.012283	0.012983	0.012749	0.013267	0.014130	0.015441	0.014597
Correl. Coeff. Act.vs.pred.	0.0361	0.0343	0.0861	0.0633	0.0724	0.0697	0.1206	0.0290	0.1205
t-value	1.76	1.67	4.20	3.09	3.53	3.40	5.88	1.41	5.88
Theil's U	0.998	1.099	0.978	1.186	0.808	1.036	0.896	0.659	0.924
Mean	0.001993	0.001930	0.001828	0.001361	0.001833	0.001443	0.001638	0.001213	0.001663
Std. Dev.	0.016105	0.015072	0.015541	0.014837	0.015234	0.014452	0.015710	0.013633	0.015565
Correct Sign Proportion	**62.12%**	60.56%	60.86%	57.08%	59.60%	55.31%	55.44%	53.59%	54.81%
Correl. Coef. sign prediction	0.2350	0.2096	0.2146	0.1464	0.1909	0.1108	0.1145	0.0909	0.1034
t-value	11.47	10.23	10.47	7.14	9.31	5.41	5.59	4.44	5.04
End wealth predict. (EW)	**11,414**	9,825	7,705	2,542	7,796	3,086	4,907	1,788	5,209
EW(pred) vs. EW(b&h)	411.70%	340.47%	245.41%	13.95%	249.52%	38.34%	120.00%	-19.83%	133.54%
Switching frequency	0.36455	0.37715	0.37295	0.39647	0.38723	0.42251	0.48635	0.44687	0.50483
Break even trans. cost	0.19%	0.17%	0.14%	0.01%	0.14%	0.03%	0.07%	0.00%	0.07%

nn = 100 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.991062	0.994211	1.006336	1.005137	1.013771	1.004189	1.006551	1.010763	1.016257
MAE	**0.011499**	0.011550	0.011585	0.011627	0.011777	0.011739	0.011778	0.011780	0.011871
Correl. Coeff. Act.vs.pred.	0.1465	0.1431	0.1040	0.1127	0.0851	0.1320	0.1338	0.1276	0.1100
t-value	7.15	6.98	5.07	5.50	4.15	6.44	6.53	6.23	5.37
Theil's U	1.003	1.025	0.981	1.026	0.928	0.849	**0.667**	0.809	0.807
Mean	0.001336	0.001276	0.001460	0.001632	0.001593	0.001580	0.001668	0.001545	0.001309
Std. Dev.	0.015356	0.015225	0.013347	0.014767	0.014502	0.014810	0.015138	0.014822	0.014816
Correct Sign Proportion	54.56%	53.84%	54.77%	54.51%	55.14%	54.09%	55.31%	55.14%	53.59%
Correl. Coef. sign prediction	0.0996	0.0904	0.1030	0.0983	0.1085	0.0862	0.1104	0.1089	0.0766
t-value	4.86	4.41	5.03	4.79	5.29	4.20	5.39	5.31	3.74
End wealth predict. (EW)	2,395	2,077	3,214	4,842	4,415	4,278	5,276	3,932	2,246
EW(pred) vs. EW(b&h)	7.36%	-6.87%	44.09%	117.05%	97.91%	91.78%	136.52%	76.25%	0.70%
Switching frequency	0.32507	0.37547	0.39143	0.38975	0.40151	0.42587	0.42587	0.40487	0.41075
Break even trans. cost	0.01%	0.00%	0.04%	0.08%	0.07%	0.06%	0.09%	0.06%	0.00%

nn = 200 / m =	2	3	4	5	6	7	8	9	10
NRMSE	**0.990298**	0.991995	1.002701	1.002685	1.008256	1.001063	1.006366	1.002440	1.005822
MAE	0.011549	0.011551	0.011584	0.011621	0.011702	0.011699	0.011761	0.011705	0.011775
Correl. Coeff. Act.vs.pred.	0.1516	**0.1534**	0.1158	0.1207	0.1014	0.1365	0.1203	0.1445	0.1363
t-value	7.40	7.49	5.65	5.89	4.95	6.66	5.87	7.05	6.65
Theil's U	1.050	1.042	1.020	1.034	0.970	0.878	0.734	0.786	0.724
Mean	0.001217	0.001132	0.001487	0.001733	0.001705	0.001511	0.001419	0.001651	0.001705
Std. Dev.	0.015070	0.014721	0.013471	0.014604	0.014821	0.014900	0.015258	0.014939	0.014987
Correct Sign Proportion	52.79%	53.38%	54.98%	55.06%	55.14%	54.18%	53.72%	54.60%	54.18%
Correl. Coef. sign prediction	0.0725	0.0893	0.1116	0.1139	0.1116	0.0904	0.0805	0.1003	0.0924
t-value	3.54	4.36	5.44	5.56	5.45	4.41	3.93	4.89	4.51
End wealth predict. (EW)	1,806	1,476	3,433	6,150	5,755	3,627	2,918	5,057	5,756
EW(pred) vs. EW(b&h)	-19.02%	-33.81%	53.88%	175.71%	158.01%	62.59%	30.80%	126.71%	158.05%
Switching frequency	0.29903	0.32927	0.36959	0.37631	0.38555	0.38219	0.41243	0.39899	0.38975
Break even trans. cost	0.00%	0.00%	0.05%	0.11%	0.10%	0.05%	0.03%	0.09%	0.10%

Non Linear predictions (Simplex method)

m =	2	3	4	5	6	7	8	9	10
NRMSE	1.008062	1.003102	0.994429	**0.985702**	0.996228	0.991566	0.993676	0.993727	1.016977
MAE	0.011774	0.011707	0.011717	**0.011527**	0.011744	0.011529	0.011591	0.011713	0.012331
Correl. Coeff. Act.vs.pred.	0.1538	0.1452	0.1554	**0.1772**	0.1308	0.1316	0.1257	0.1383	0.0883
t-value	7.51	7.09	7.58	8.65	6.38	6.42	6.13	6.75	4.31
Theil's U	1.561	1.256	1.220	1.399	1.280	1.013	1.143	**0.721**	1.597
Mean	0.001656	0.001800	0.001751	0.001955	0.001458	0.001469	0.001527	0.001402	0.000600
Std. Dev.	0.015791	0.015161	0.015217	0.015942	0.014558	0.016561	0.013815	0.018704	0.008571
Correct Sign Proportion	57.24%	58.92%	56.28%	**57.87%**	53.55%	56.36%	55.27%	54.64%	48.76%
Correl. Coef. sign prediction	0.1441	0.1769	0.1378	0.1652	0.0972	0.1083	0.1129	0.0478	0.0577
t-value	7.03	8.63	6.73	8.06	4.74	5.28	5.51	2.33	2.82
End wealth predict. (EW)	5,125	7,212	6,426	**10,409**	3,197	3,283	3,769	2,802	417
EW(pred) vs. EW(b&h)	129.74%	223.33%	188.09%	366.62%	43.34%	47.20%	68.98%	25.60%	-81.30%
Switching frequency	0.41075	0.39647	0.36791	0.39647	0.35699	0.37967	0.41243	0.14532	0.16128
Break even trans. cost	0.09%	0.12%	0.12%	0.16%	0.04%	0.04%	0.05%	0.07%	0.00%

Table 8.4 : ASE Prediction Results (Library : 4- year, Prediction period : 9- year)

Linear Predictions

Prediction model	Random Walk	MA(20)	Linear model
NRMSE	1.218109	1.012357	**0.973166**
MAE	0.014224	0.012689	**0.011837**
Correl. Coeff. Act.vs.pred.	0.2581	0.1019	**0.2680**
t-value	11.91	4.70	12.37
Theil's U	1.000	0.889	0.852
Mean	0.003061	0.002026	0.003175
Std. Dev.	0.016077	0.014560	0.016344
Correct Sign Proportion	**64.20%**	57.53%	63.87%
Correl. Coef. sign prediction	0.2796	0.1414	0.2672
t-value	12.91	6.53	12.33
End wealth predict. (EW)	66,959	7,438	**85,227**
EW(pred) vs. EW(b&h)	3428.12%	291.92%	4390.68%
Switching frequency	0.35758	0.06007	0.39794
Break even trans. cost	**0.47%**	**1.06%**	**0.45%**

Market Index (Buy & hold strategy)

Mean	0.0013834
Std. Dev.	0.0206206
End Wealth	1,898

Non Linear predictions (Piecewise approx. method)

nn = 20 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.281702	1.227001	1.270202	1.316234	1.312840	1.329761	1.458281	1.577445	1.642267
MAE	0.013840	0.014080	0.014459	0.015306	0.015743	0.016270	0.017878	0.019298	0.019673
Correl. Coeff. Act.vs.pred.	-0.0069	0.0016	0.0128	0.0295	0.0806	0.1303	0.1227	0.1458	0.1672
t-value	-0.32	0.07	0.59	1.36	3.72	6.01	5.66	6.73	7.72
Theil's U	1.256	1.171	1.019	0.780	1.206	1.146	0.944	2.601	1.590
Mean	0.001815	0.001502	0.001385	0.001609	0.001568	0.001550	0.001505	0.001637	0.001667
Std. Dev.	0.016769	0.015985	0.016298	0.015748	0.016503	0.017596	0.017144	0.017231	0.016941
Correct Sign Proportion	59.69%	59.13%	59.08%	57.16%	55.89%	55.80%	53.82%	54.48%	54.67%
Correl. Coef. sign prediction	0.1878	0.1805	0.1785	0.1401	0.1167	0.1123	0.0764	0.0896	0.0929
t-value	8.67	8.33	8.24	6.47	5.39	5.18	3.53	4.13	4.29
End wealth predict. (EW)	4,746	2,443	4,234	3,065	2,811	2,702	2,460	3,255	3,471
EW(pred) vs. EW(b&h)	150.06%	28.72%	123.09%	61.50%	48.11%	42.35%	29.60%	71.50%	82.88%
Switching frequency	0.38855	0.40732	0.40638	0.46269	0.46645	0.45800	0.46551	0.45237	0.46926
Break even trans. cost	0.11%	0.03%	0.09%	0.05%	0.04%	0.04%	0.03%	0.06%	0.06%

nn = 50 / m =	2	3	4	5	6	7	8	9	10
	1.135590	1.117598	1.117413	1.151989	1.157924	1.146968	1.221510	1.327447	1.246304
MAE	0.012939	0.013109	0.013160	0.013719	0.013664	0.013959	0.014608	0.015968	0.015389
Correl. Coeff. Act.vs.pred.	0.0176	0.0306	0.0649	0.0540	0.0624	0.0624	0.1029	0.0460	0.1361
t-value	0.81	1.41	2.99	2.49	2.88	2.88	4.75	2.13	6.28
Theil's U	1.048	1.121	0.839	1.240	0.918	0.912	1.065	0.881	1.076
Mean	0.001841	0.001878	0.001716	0.001410	0.001798	0.001480	0.001784	0.001308	0.001710
Std. Dev.	0.016853	0.015601	0.016472	0.015692	0.016234	0.016232	0.017017	0.015082	0.016398
Correct Sign Proportion	**61.15%**	60.58%	59.31%	56.92%	59.55%	56.50%	56.12%	54.06%	55.84%
Correl. Coef. sign prediction	0.2148	0.2092	0.1810	0.1398	0.1867	0.1327	0.1253	0.0956	0.1215
t-value	9.92	9.66	8.35	6.46	8.62	6.12	5.78	4.41	5.61
End wealth predict. (EW)	5,022	5,436	3,852	2,006	4,577	2,332	4,451	1,616	3,798
EW(pred) vs. EW(b&h)	164.62%	186.40%	102.99%	5.72%	141.16%	22.88%	134.55%	-14.86%	100.12%
Switching frequency	0.37823	0.39418	0.39137	0.41389	0.42234	0.42891	0.48334	0.43548	0.49179
Break even trans. cost	0.12%	0.13%	0.09%	0.01%	0.10%	0.02%	0.08%	0.00%	0.07%

nn = 100 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.987727	0.990147	0.996908	0.993665	0.999704	0.991614	0.997505	1.003125	1.011271
MAE	**0.012370**	0.012402	0.012382	0.012392	0.012505	0.012541	0.012539	0.012599	0.012673
Correl. Coeff. Act.vs.pred.	0.1630	0.1614	0.1451	0.1641	0.1472	**0.1848**	0.1729	0.1618	0.1428
t-value	7.53	7.45	6.70	7.58	6.80	8.53	7.98	7.47	6.59
Theil's U	1.021	0.932	0.953	0.977	0.819	0.675	0.650	**0.632**	0.779
Mean	0.001614	0.001655	0.001758	0.001851	0.001853	0.001694	0.001705	0.001589	0.001581
Std. Dev.	0.016249	0.015927	0.014395	0.015564	0.015852	0.015894	0.016033	0.016363	0.016580
Correct Sign Proportion	55.61%	55.94%	56.87%	56.08%	55.56%	55.14%	56.26%	55.47%	55.51%
Correl. Coef. sign prediction	0.1133	0.1271	0.1402	0.1260	0.1119	0.1028	0.1248	0.1090	0.1094
t-value	5.23	5.87	6.47	5.81	5.16	4.74	5.76	5.03	5.05
End wealth predict. (EW)	3,097	3,381	4,206	5,122	5,153	3,676	3,758	2,936	2,888
EW(pred) vs. EW(b&h)	63.20%	78.13%	121.64%	169.90%	171.53%	93.68%	97.99%	54.68%	52.17%
Switching frequency	0.33599	0.38198	0.40357	0.40544	0.56312	0.58752	0.41671	0.40920	0.39981
Break even trans. cost	0.07%	0.07%	0.09%	0.12%	0.08%	0.07%	0.08%	0.05%	0.05%

nn = 200 / m =	2	3	4	5	6	7	8	9	10
NRMSE	**0.986082**	0.988735	0.994573	0.992802	0.996480	0.991924	0.997021	0.993483	0.998268
MAE	0.012406	0.012427	0.012398	0.012418	0.012456	0.012488	0.012522	0.012507	0.012544
Correl. Coeff. Act.vs.pred.	0.1729	0.1687	0.1512	0.1621	0.1509	0.1750	0.1564	0.1802	0.1667
t-value	7.98	7.79	6.98	7.48	6.97	8.08	7.22	8.32	7.70
Theil's U	1.062	1.010	0.986	1.028	0.875	0.752	0.728	0.647	0.753
Mean	0.001514	0.001589	0.001724	0.001847	0.001888	0.001570	0.001600	0.001755	0.001743
Std. Dev.	0.016046	0.015786	0.015683	0.015540	0.015709	0.015835	0.016438	0.016310	0.016313
Correct Sign Proportion	55.00%	54.58%	56.12%	55.84%	56.03%	54.95%	54.81%	55.51%	55.75%
Correl. Coef. sign prediction	0.1092	0.1077	0.1306	0.1256	0.1249	0.1020	0.0980	0.1127	0.1170
t-value	5.04	4.97	6.03	5.80	5.77	4.71	4.52	5.20	5.40
End wealth predict. (EW)	2,507	2,938	3,915	5,084	**5,550**	2,822	3,009	4,180	4,080
EW(pred) vs. EW(b&h)	32.08%	54.78%	106.28%	167.88%	192.41%	48.67%	58.53%	120.26%	114.96%
Switching frequency	0.29188	0.34162	0.37353	0.39887	0.39230	0.38667	0.39794	0.40544	0.38949
Break even trans. cost	0.05%	0.06%	0.09%	0.12%	0.13%	0.05%	0.05%	0.09%	0.09%

Non Linear predictions (Simplex method)

m =	2	3	4	5	6	7	8	9	10
NRMSE	1.012729	1.002050	0.995856	**0.986881**	0.995017	0.989027	0.990695	0.989729	1.014574
MAE	0.012710	0.012544	0.012591	0.012438	0.012635	**0.012402**	0.012465	0.012572	0.013202
Correl. Coeff. Act.vs.pred.	0.1333	0.1411	0.1509	**0.1728**	0.1421	0.1484	0.1496	0.1741	0.1190
t-value	6.15	6.52	6.97	7.98	6.56	6.85	6.91	8.04	5.49
Theil's U	1.289	1.138	1.141	1.074	1.268	0.910	1.024	**0.752**	1.638
Mean	0.001595	0.001750	0.001705	0.001943	0.001478	0.001834	0.001862	0.001579	0.000585
Std. Dev.	0.016340	0.015438	0.014391	0.014931	0.013310	0.017145	0.016051	0.019880	0.008679
Correct Sign Proportion	57.63%	59.74%	57.48%	**58.52%**	54.48%	57.63%	56.73%	55.04%	47.91%
Correl. Coef. sign prediction	0.1481	0.1917	0.1619	0.1763	0.1184	0.1357	0.1409	0.0663	0.0355
t-value	6.84	8.85	7.47	8.14	5.47	6.26	6.51	3.06	1.64
End wealth predict. (EW)	2,977	4,139	3,762	**6,231**	2,320	4,941	5,244	2,877	347
EW(pred) vs. EW(b&h)	56.86%	118.10%	98.21%	228.34%	22.26%	160.36%	176.33%	51.59%	-81.70%
Switching frequency	0.40920	0.39512	0.40357	0.39043	0.36321	0.36321	0.39418	0.11919	0.14172
Break even trans. cost	0.05%	0.09%	0.08%	0.14%	0.03%	0.12%	0.12%	0.16%	0.00%

Table 8.5 : ASE Prediction Results (Library : 5- year, Prediction period : 8- year)

Linear Predictions

Prediction model	Random Walk	MA(20)	Linear model
NRMSE	1.220800	1.012463	0.974115
MAE	0.015373	0.013639	**0.012770**
Correl. Coeff. Act.vs.pred.	0.2548	0.1013	**0.2664**
t-value	11.05	4.40	11.55
Theil's U	1.000	0.886	0.915
Mean	0.003202	0.002063	0.003347
Std. Dev.	0.016906	0.015225	0.017102
Correct Sign Proportion	63.48%	56.51%	**63.48%**
Correl. Coef. sign prediction	0.2678	0.1266	0.2656
t-value	11.61	5.49	11.52
End wealth predict. (EW)	40,608	4,802	**53,314**
EW(pred) vs. EW(b&h)	2888.39%	253.40%	3823.42%
Switching frequency	0.36470	0.06273	0.40510
Break even trans. cost	**0.49%**	**1.07%**	**0.48%**

Market Index (Buy & hold strategy)	
Mean	0.0013896
Std. Dev.	0.0217113
End Wealth	1,359

Non Linear predictions (Piecewise approx. method)

nn = 20 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.037239	1.100031	1.171236	1.219419	1.245218	1.325969	1.500950	1.533246	1.739415
MAE	0.013457	0.014048	0.014489	0.015485	0.015875	0.017290	0.018707	0.019567	0.021697
Correl. Coeff. Act.vs.pred.	0.1515	0.1096	0.0957	0.1233	0.1925	0.1449	0.1092	0.1012	0.0767
t-value	6.57	4.75	4.15	5.35	8.35	6.28	4.74	4.39	3.33
Theil's U	1.192	1.837	1.309	2.233	1.214	2.522	2.108	1.896	1.304
Mean	0.002500	0.001913	0.002097	0.001866	0.002051	0.001613	0.001423	0.001651	0.001165
Std. Dev.	0.015056	0.014973	0.014485	0.013225	0.014596	0.015450	0.015135	0.015750	0.014614
Correct Sign Proportion	61.40%	59.17%	60.29%	58.90%	57.47%	55.50%	54.49%	55.24%	53.91%
Correl. Coef. sign prediction	0.2269	0.1845	0.2090	0.1803	0.1491	0.1096	0.0901	0.1047	0.0786
t-value	9.84	8.00	9.06	7.82	6.47	4.75	3.91	4.54	3.41
End wealth predict. (EW)	10,913	3,625	5,123	3,324	4,701	2,067	1,446	2,220	891
EW(pred) vs. EW(b&h)	703.11%	166.74%	277.03%	144.58%	245.94%	52.12%	6.38%	63.40%	-34.44%
Switching frequency	0.39979	0.41680	0.42531	0.44125	0.45189	0.47953	0.46890	0.45614	0.46146
Break even trans. cost	0.28%	0.13%	0.17%	0.11%	0.15%	0.05%	0.01%	0.06%	0.00%

nn = 50 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.993317	1.024245	1.076951	1.093820	1.081605	1.133209	1.166479	1.258831	1.164468
MAE	**0.013088**	0.013372	0.013703	0.014247	0.014127	0.014823	0.015323	0.016363	0.015794
Correl. Coeff. Act.vs.pred.	0.2030	0.1662	0.1105	0.1016	0.1484	0.1342	0.1360	0.0891	0.1525
t-value	8.81	7.21	4.79	4.41	6.44	5.82	5.90	3.86	6.61
Theil's U	1.091	1.211	0.847	1.385	1.030	1.041	0.787	0.987	0.758
Mean	0.002856	0.002405	0.002261	0.001851	0.002154	0.001867	0.001841	0.001416	0.001922
Std. Dev.	0.015691	0.015376	0.014382	0.014448	0.015743	0.016090	0.016193	0.014231	0.016591
Correct Sign Proportion	**62.52%**	60.98%	60.66%	57.52%	60.34%	57.10%	56.78%	54.86%	56.83%
Correl. Coef. sign prediction	0.2480	0.2184	0.2142	0.1536	0.2049	0.1424	0.1352	0.1041	0.1373
t-value	10.76	9.47	9.29	6.66	8.88	6.17	5.87	4.51	5.95
End wealth predict. (EW)	**21,232**	9,126	6,960	3,232	5,695	3,330	3,171	1,427	3,692
EW(pred) vs. EW(b&h)	1462.48%	571.55%	412.21%	137.82%	319.07%	145.07%	133.39%	5.03%	171.73%
Switching frequency	0.40191	0.39872	0.42212	0.44125	0.44338	0.46784	0.48910	0.43275	0.47634
Break even trans. cost	**0.36%**	0.25%	0.21%	0.11%	0.17%	0.10%	0.09%	0.01%	0.11%

nn = 100 / m =	2	3	4	5	6	7	8	9	10
NRMSE	**0.969117**	0.973681	0.986968	0.990765	0.986796	0.986316	0.988921	0.995353	0.998072
MAE	0.013119	0.013106	0.013202	0.013262	0.013355	0.013404	0.013494	0.013536	0.013521
Correl. Coeff. Act.vs.pred.	**0.2466**	0.2322	0.1990	0.1951	0.2121	0.2213	0.2189	0.2042	0.1982
t-value	10.70	10.07	8.63	8.46	9.20	9.60	9.49	8.86	8.60
Theil's U	1.017	1.026	0.964	1.033	0.913	0.694	0.663	0.893	0.921
Mean	0.002236	0.002194	0.002378	0.002170	0.002218	0.002117	0.002125	0.001904	0.001788
Std. Dev.	0.015834	0.016458	0.015156	0.015280	0.017312	0.016873	0.016793	0.016588	0.016776
Correct Sign Proportion	56.67%	57.74%	58.37%	57.84%	57.36%	56.62%	57.58%	56.46%	56.25%
Correl. Coef. sign prediction	0.1319	0.1559	0.1683	0.1561	0.1443	0.1296	0.1483	0.1260	0.1202
t-value	5.72	6.76	7.30	6.77	6.26	5.62	6.43	5.47	5.21
End wealth predict. (EW)	6,648	6,140	8,670	5,870	6,426	5,318	5,403	3,564	2,868
EW(pred) vs. EW(b&h)	389.22%	351.84%	538.05%	331.98%	372.86%	291.33%	297.62%	162.27%	111.06%
Switching frequency	0.32111	0.36683	0.39979	0.41148	0.39022	0.43169	0.41786	0.42105	0.42424
Break even trans. cost	0.26%	0.22%	0.25%	0.19%	0.21%	0.17%	0.18%	0.12%	0.09%

nn = 200 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.970609	0.974070	0.984135	0.986424	0.984297	0.982936	0.984627	0.982519	0.987333
MAE	0.013143	0.013140	0.013196	0.013217	0.013289	0.013318	0.013323	0.013308	0.013332
Correl. Coeff. Act.vs.pred.	0.2411	0.2306	0.2034	0.2004	0.2107	0.2188	0.2124	0.2284	0.2139
t-value	10.46	10.00	8.82	8.69	9.14	9.49	9.21	9.90	9.28
Theil's U	1.067	1.011	0.983	1.038	0.934	0.757	0.699	**0.643**	0.730
Mean	0.002248	0.002180	0.002255	0.002228	0.002195	0.002234	0.002165	0.002108	0.002088
Std. Dev.	0.015775	0.016628	0.014769	0.015010	0.017242	0.016952	0.016934	0.016842	0.016530
Correct Sign Proportion	56.67%	56.78%	58.53%	58.43%	57.42%	56.94%	57.04%	56.94%	57.20%
Correl. Coef. sign prediction	0.1360	0.1409	0.1738	0.1708	0.1473	0.1380	0.1382	0.1360	0.1411
t-value	5.90	6.11	7.54	7.41	6.39	5.98	5.99	5.90	6.12
End wealth predict. (EW)	6,803	5,980	6,885	6,549	6,159	6,628	5,822	5,231	5,032
EW(pred) vs. EW(b&h)	400.62%	340.05%	406.69%	381.95%	353.24%	387.75%	328.43%	284.96%	270.32%
Switching frequency	0.31047	0.33174	0.38490	0.38915	0.38384	0.38915	0.39022	0.40829	0.38490
Break even trans. cost	0.28%	0.24%	0.22%	0.22%	0.21%	0.22%	0.20%	0.18%	0.18%

Non Linear predictions (Simplex method)

m =	2	3	4	5	6	7	8	9	10
NRMSE	0.998130	0.988994	0.983387	**0.981482**	0.985615	0.983793	0.985148	0.987259	1.009733
MAE	0.013594	0.013301	0.013277	**0.013234**	0.013372	0.013327	0.013310	0.013522	0.013999
Correl. Coeff. Act.vs.pred.	0.1956	0.1991	**0.2065**	0.2001	0.1909	0.1800	0.1835	0.1940	0.1510
t-value	8.48	8.64	8.96	8.68	8.28	7.81	7.96	8.41	6.55
Theil's U	1.289	1.091	1.081	1.254	1.399	1.069	1.141	**0.736**	1.698
Mean	0.002212	0.002070	0.002086	0.002243	0.002086	0.001804	0.001848	0.001705	0.000730
Std. Dev.	0.015422	0.015726	0.014644	0.014334	0.013731	0.018221	0.016709	0.021129	0.007930
Correct Sign Proportion	58.96%	**60.29%**	59.38%	59.38%	57.31%	56.30%	56.78%	53.85%	49.18%
Correl. Coef. sign prediction	0.1780	0.2034	0.1986	0.1906	0.1649	0.1162	0.1387	0.0649	0.0376
t-value	7.72	8.82	8.62	8.27	7.15	5.04	6.02	2.81	1.63
End wealth predict. (EW)	6,356	4,873	5,020	**6,739**	5,018	2,959	3,212	2,457	394
EW(pred) vs. EW(b&h)	367.77%	258.59%	269.45%	395.95%	269.29%	117.74%	136.40%	80.79%	-70.99%
Switching frequency	0.42637	0.40404	0.38171	0.39022	0.34662	0.36683	0.38596	0.13184	0.13291
Break even trans. cost	0.19%	0.17%	0.18%	0.22%	0.20%	0.11%	0.12%	0.24%	0.00%

The linear model and the random walk model exhibit the best correlation coefficient indicators and this time the improvement over the performance of the nonlinear models exceeds 30% for 1-year library length. However, the performance of the nonlinear models appears to be an increasing function with the library length, an expected characteristic especially when a chaotic component is present.
Finally, Theil's U statistic performs better for the PW nonlinear model.

As already discussed, statistical significance of predictability does not necessarily translates into profit making possibilities. Assessing the economic significance of the predictions in Tables 8.1 to 8.5 we can see that all the switching portfolios outperform the market portfolio (buy and hold) by far in a mean-variance sense and this is also a statistically significant result.
Specifically, the linear RW and AR models and the nonlinear PW, exhibit correct sign indicator ranging from 61% to 65% and a statistically significant correlation coefficient between the realised and predicted sign. The same is true for the MA and the SX models; yet, the sign indicator is lower, ranging between 55%-58%.
In terms of end wealth, the dominance of the RW and the AR models over the nonlinear specifications is complete. The end wealth of these two linear switching portfolios exceeds the market portfolio end wealth by 2888% to 8485% depending on the prediction period scenario. The performance of the nonlinear switching portfolios is also high compared to the market portfolio, which is outperformed by 228% to 1486%.
In all cases, the linear switching portfolios have a higher mean and a lower variance than the mean and variance corresponding to the best estimates[6] of the nonlinear portfolios.
In terms of the break even transaction cost, the MA portfolio is the only one to exceed the average effective transaction cost for all periods. It should be noticed that, in the case of the MA portfolio, this is due to the very low switching frequency (less than 5% on average), minimising the overall transaction cost. Linear portfolios give, on average, higher break even transaction cost than the nonlinear ones. This cost is close to the market's effective transaction cost.

---

[6] In the case of the nonlinear portfolios, multiple prediction sets are produced corresponding to different parameter values. For our comparisons we use the best estimates obtained by each model.

Among the nonlinear portfolios the PW model performs better than the SX model, nevertheless, both models give a break even transaction cost much lower than the effective one.

The above results show a remarkable stability and do not seem to be affected by the library length or the prediction period employed. The statistical measures across periods are very similar for each model.

In the case of the PW model, statistical measures such as the NRMSE, MAE and Theil's U are less correlated to the end wealth value across periods. Yet, they show an almost perfect correlation in the case of the SX model, where the best NRMSE, MAE and CC values, in all prediction periods, coincide with the best end wealth value for m=5. However, it should be noted that the best performance of the PW model in terms of economic results (end wealth), is obtained for low embedding dimensions (m=2-4) and a moderate number of nearest neighbours (NN=50). Accordingly, for the SX model, the best results are obtained with m=5. These findings, combined with the fact that the best values of the correlation coefficient between realised and predicted values are also obtained for small embedding dimensions, could be considered as an indication of a masked chaotic component in the ASE series. This is an indirect indication since the analysis above does not directly serve as a test for chaoticity.

The corresponding results for the **LSE returns** are presented in Tables 8.6 to 8.10. This time, predictability indications are weaker. The best value of the correlation coefficient measure is 13% and statistically significant, yet this is only half of the respective value found for the ASE returns.

In terms of statistical measures, the nonlinear PW model outperforms all other models across all periods. In terms of NRMSE, it is the only model exhibiting a lower value than 1; nevertheless, its superiority over the linear models in terms of both the NRMSE and MAE measures, is marginal.

In terms of the CC measure, improvement over the random walk model is marginal too, but exceeds 50% when compared to the CC of the AR(1) model.

Assessing the economic significance of predictions in Tables 8.6 to 8.10, we observe that the switching portfolios based on predictions from the RW, the AR and the PW

Table 8.6 : LSE Prediction Results (Library : 1- year, Prediction period : 12- year)

Linear Predictions

Prediction model	Random Walk	MA(20)	Linear model
NRMSE	1.321275	1.021548	1.020649
MAE	0.008365	0.006463	0.006320
Correl. Coeff. Act.vs.pred.	0.1271	0.0675	0.0086
t-value	7.06	3.75	0.48
Theil's U	1.000	1.066	0.999
Mean	0.000825	0.000637	0.000776
Std. Dev.	0.005903	0.005988	0.006933
Correct Sign Proportion	56.58%	54.99%	57.58%
Correl. Coef. sign prediction	0.1382	0.0953	0.1512
t-value	7.67	5.29	8.40
End wealth predict. (EW)	1,273	713	1,093
EW(pred) vs. EW(b&h)	54.65%	-13.38%	32.75%
Switching frequency	0.45658	0.07485	0.27641
Break even trans. cost	0.031%	0.000%	0.034%

Market Index (Buy & hold strategy)	
Mean	0.0006838
Std. Dev.	0.0086315
End Wealth	823

Non Linear predictions (Piecewise approx. method)

nn = 20 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.105955	1.122071	1.192805	1.177150	1.152144	1.210944	1.238810	1.351801	1.390887
MAE	0.006667	0.006816	0.007051	0.007296	0.007261	0.007545	0.007899	0.008106	0.008481
Correl. Coeff. Act.vs.pred.	0.0207	0.0665	-0.0363	0.0230	0.0618	0.0642	0.0710	-0.0178	0.0194
t-value	1.15	3.70	-2.01	1.28	3.43	3.57	3.94	-0.99	1.08
Theil's U	1.020	1.167	0.948	1.051	1.071	1.072	1.253	1.349	1.158
Mean	0.000690	0.000640	0.000619	0.000477	0.000500	0.000484	0.000470	0.000352	0.000413
Std. Dev.	0.006964	0.006894	0.006811	0.006631	0.006346	0.005969	0.005958	0.006332	0.005951
Correct Sign Proportion	57.00%	55.77%	56.58%	54.54%	54.05%	53.69%	53.76%	53.27%	53.86%
Correl. Coef. sign prediction	0.1390	0.1122	0.1318	0.0916	0.0850	0.0766	0.0822	0.0714	0.0907
t-value	7.72	6.24	7.32	5.09	4.72	4.26	4.57	3.96	5.04
End wealth predict. (EW)	838	720	674	435	467	445	426	296	357
EW(pred) vs. EW(b&h)	1.79%	-12.56%	-18.13%	-47.14%	-43.26%	-45.97%	-48.29%	-64.09%	-56.59%
Switching frequency	0.46954	0.48218	0.49028	0.51620	0.52430	0.51977	0.50065	0.50000	0.48736
Break even trans. cost	0.001%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%

nn = 50 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.011458	1.009683	1.025020	1.040756	1.030483	1.039721	1.079172	1.084934	1.083470
MAE	0.006344	0.006384	0.006434	0.006571	0.006541	0.006580	0.006874	0.006910	0.006931
Correl. Coeff. Act.vs.pred.	0.0399	0.0834	0.0279	-0.0007	0.0377	0.0436	-0.0067	-0.0057	0.0151
t-value	2.22	4.63	1.55	-0.04	2.09	2.42	-0.37	-0.31	0.84
Theil's U	1.005	1.006	0.972	1.048	1.030	0.982	1.060	1.032	1.037
Mean	0.000679	0.000612	0.000687	0.000531	0.000584	0.000595	0.000383	0.000344	0.000357
Std. Dev.	0.007504	0.007285	0.007119	0.006926	0.006923	0.006738	0.006875	0.006707	0.006597
Correct Sign Proportion	57.42%	56.87%	57.91%	55.64%	56.16%	56.45%	54.76%	55.12%	54.37%
Correl. Coef. sign prediction	0.1508	0.1346	0.1606	0.1042	0.1173	0.1247	0.0868	0.0977	0.0798
t-value	8.38	7.48	8.92	5.79	6.52	6.93	4.82	5.43	4.43
End wealth predict. (EW)	811	660	832	514	605	627	326	289	301
EW(pred) vs. EW(b&h)	-1.46%	-19.86%	1.04%	-37.54%	-26.55%	-23.85%	-60.43%	-64.95%	-63.47%
Switching frequency	0.30817	0.33733	0.31821	0.38464	0.40117	0.38756	0.33441	0.32469	0.32923
Break even trans. cost	0.000%	0.000%	0.001%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%

nn = 100 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.003638	1.006508	1.012496	1.020672	1.024815	1.030653	1.040701	1.053496	1.051719
MAE	0.006338	0.006359	0.006362	0.006434	0.006469	0.006487	0.006558	0.006663	0.006637
Correl. Coeff. Act.vs.pred.	0.0779	0.0915	0.0513	0.0341	0.0385	0.0351	0.0180	0.0014	0.0376
t-value	4.33	5.08	2.85	1.90	2.14	1.95	1.00	0.08	2.09
Theil's U	1.012	0.989	0.984	1.004	1.024	1.000	1.015	1.031	1.021
Mean	0.000683	0.000660	0.000696	0.000673	0.000576	0.000665	0.000526	0.000486	0.000597
Std. Dev.	0.007311	0.007163	0.007294	0.007005	0.006874	0.006731	0.006803	0.006673	0.006681
Correct Sign Proportion	57.36%	57.32%	58.00%	57.19%	55.96%	57.03%	56.45%	55.28%	56.58%
Correl. Coef. sign prediction	0.1536	0.1469	0.1639	0.1420	0.1129	0.1376	0.1251	0.0986	0.1297
t-value	8.53	8.16	9.11	7.89	6.27	7.64	6.95	5.48	7.20
End wealth predict. (EW)	821	765	856	796	591	778	506	448	630
EW(pred) vs. EW(b&h)	-0.22%	-7.05%	3.96%	-3.35%	-28.22%	-5.49%	-38.50%	-45.58%	-23.49%
Switching frequency	0.27965	0.27965	0.29715	0.34251	0.30622	0.36973	0.36261	0.30428	0.40797
Break even trans. cost	0.000%	0.000%	0.004%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%

nn = 200 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.992101	0.991597	0.993861	0.993410	0.996414	1.007488	1.005908	1.014125	1.009813
MAE	0.006278	0.006289	0.006291	0.006297	0.006308	0.006370	0.006351	0.006421	0.006390
Correl. Coeff. Act.vs.pred.	0.1304	0.1304	0.1143	0.1166	0.1034	0.0780	0.0739	0.0580	0.0692
t-value	7.25	7.24	6.35	6.48	5.75	4.33	4.10	3.22	3.85
Theil's U	1.004	1.003	1.001	1.002	1.017	0.989	0.987	0.994	0.977
Mean	0.000873	0.000940	0.000838	0.000904	0.000806	0.000699	0.000703	0.000655	0.000725
Std. Dev.	0.006475	0.006570	0.006584	0.006625	0.006428	0.006368	0.006453	0.006147	0.006286
Correct Sign Proportion	57.87%	58.39%	57.71%	58.33%	56.90%	56.32%	56.74%	55.09%	56.29%
Correl. Coef. sign prediction	0.1578	0.1698	0.1539	0.1687	0.1359	0.1276	0.1357	0.1030	0.1266
t-value	8.77	9.43	8.55	9.37	7.55	7.09	7.54	5.72	7.03
End wealth predict. (EW)	1,474	1,815	1,325	1,621	1,200	864	874	753	936
EW(pred) vs. EW(b&h)	79.07%	120.43%	60.88%	96.87%	45.80%	4.88%	6.18%	-8.49%	13.65%
Switching frequency	0.35872	0.32372	0.28224	0.30233	0.32178	0.47861	0.46176	0.48218	0.52268
Break even trans. cost	0.053%	0.080%	0.055%	0.073%	0.038%	0.003%	0.004%	0.000%	0.008%

Non Linear predictions (Simplex method)

m =	2	3	4	5	6	7	8	9	10
NRMSE	1.148514	1.112155	1.099786	1.059271	1.046288	1.095834	1.093543	1.052750	**1.022345**
MAE	0.007428	0.007218	0.007155	0.006760	0.006642	0.007164	0.007179	0.006817	**0.006450**
Correl. Coeff. Act.vs.pred.	0.0159	0.0315	0.0242	0.0201	0.0134	-0.0039	-0.0026	0.0224	**0.0341**
t-value	0.88	2.14	1.35	1.11	0.74	-0.22	-0.14	1.24	1.89
Theil's U	1.162	1.070	1.086	1.081	1.108	1.089	1.110	1.103	**1.020**
Mean	0.000399	0.000413	0.000329	0.000399	0.000465	0.000194	0.000081	0.000340	0.000651
Std. Dev.	0.006326	0.006624	0.006247	0.006859	0.007044	0.004980	0.004422	0.004971	0.007355
Correct Sign Proportion	53.63%	54.12%	52.04%	54.50%	54.96%	48.28%	47.02%	50.19%	**58.23%**
Correl. Coef. sign prediction	0.0754	0.0843	0.0586	0.0916	0.0972	0.0055	-0.0164	0.0479	0.1660
t-value	4.19	4.68	3.25	5.09	5.40	0.31	-0.91	2.66	9.22
End wealth predict. (EW)	342	357	276	343	420	182	129	285	**744**
EW(pred) vs. EW(b&h)	-58.48%	-56.63%	-66.52%	-58.40%	-49.03%	-77.88%	-84.38%	-65.38%	-9.64%
Switching frequency	0.50713	0.52106	0.45301	0.48250	0.47634	0.31270	0.29618	0.36585	0.40311
Break even trans. cost	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%

Table 8.7 : LSE Prediction Results (Library : 2- year, Prediction period : 11- year)

Linear Predictions

Prediction model	Random Walk	MA(20)	Linear model
NRMSE	1.316529	1.020216	1.008180
MAE	0.008060	0.006236	0.006066
Correl. Coeff. Act.vs.pred.	0.1334	0.0781	0.0438
t-value	7.09	4.15	2.33
Theil's U	1.000	1.069	1.000
Mean	0.000810	0.000677	0.000824
Std. Dev.	0.005846	0.005815	0.007006
Correct Sign Proportion	54.25%	52.97%	56.30%
Correl. Coef. sign prediction	0.0687	0.0235	0.0633
t-value	3.65	1.25	3.36
End wealth predict. (EW)	984	677	1,023
EW(pred) vs. EW(b&h)	56.03%	7.32%	62.27%
Switching frequency	0.45789	0.06936	0.29936
Break even trans. cost	0.034%	0.035%	0.057%

Market Index (Buy & hold strategy)

Mean	0.0006524
Std. Dev.	0.0086106
End Wealth	631

Non Linear predictions (Piecewise approx. method)

nn = 20 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.090855	1.087726	1.203567	1.178307	1.138782	1.183306	1.344464	1.444897	1.487133
MAE	0.006389	0.006550	0.006804	0.006871	0.006985	0.007282	0.007856	0.008246	0.008599
Correl. Coeff. Act.vs.pred.	0.0468	0.1286	0.0051	0.0650	0.1667	**0.1810**	0.0224	-0.0097	0.0524
t-value	2.49	6.84	0.27	3.45	8.86	9.62	1.19	-0.52	2.79
Theil's U	1.076	1.057	**0.961**	1.204	1.251	1.190	1.320	1.309	1.451
Mean	0.000656	0.000579	0.000530	0.000514	0.000584	0.000559	0.000410	0.000393	0.000482
Std. Dev.	0.006906	0.006851	0.006730	0.006430	0.005910	0.005992	0.006117	0.006574	0.005887
Correct Sign Proportion	54.42%	53.01%	53.54%	52.44%	52.26%	51.49%	50.64%	51.42%	51.24%
Correl. Coef. sign prediction	0.0483	0.0255	0.0363	0.0219	0.0313	0.0146	0.0038	0.0148	0.0221
t-value	2.57	1.35	1.93	1.16	1.67	0.78	0.20	0.79	1.17
End wealth predict. (EW)	637	513	446	426	520	485	318	303	389
EW(pred) vs. EW(b&h)	0.93%	-18.70%	-29.29%	-32.40%	-17.61%	-23.15%	-49.57%	-51.89%	-38.25%
Switching frequency	0.45789	0.47983	0.49328	0.47488	0.48620	0.52795	0.50248	0.47912	0.50177
Break even trans. cost	0.001%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%

nn = 50 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.017917	1.013790	1.040625	1.056749	1.024676	1.033161	1.051957	1.074519	1.086948
MAE	0.006117	0.006126	0.006226	0.006293	0.006272	0.006313	0.006518	0.006581	0.006620
Correl. Coeff. Act.vs.pred.	0.0200	0.1112	0.0654	0.0202	0.0872	0.0978	0.0982	0.0442	0.0590
t-value	1.06	5.91	3.48	1.07	4.63	5.20	5.22	2.35	3.13
Theil's U	0.996	1.028	0.990	1.010	0.984	0.964	1.018	0.999	1.095
Mean	0.000699	0.000650	0.000594	0.000634	0.000668	0.000597	0.000578	0.000380	0.000510
Std. Dev.	0.007484	0.007497	0.007197	0.007219	0.006693	0.006625	0.006304	0.006384	0.005958
Correct Sign Proportion	55.02%	55.38%	54.67%	54.88%	54.21%	54.10%	53.68%	52.16%	53.26%
Correl. Coef. sign prediction	0.0336	0.0444	0.0318	0.0362	0.0278	0.0272	0.0217	-0.0008	0.0256
t-value	1.79	2.36	1.69	1.92	1.48	1.44	1.15	-0.04	1.36
End wealth predict. (EW)	719	626	536	599	659	539	511	292	422
EW(pred) vs. EW(b&h)	13.99%	-0.72%	-15.07%	-5.04%	4.41%	-14.48%	-18.99%	-53.70%	-33.02%
Switching frequency	0.36801	0.39844	0.38358	0.37438	0.39986	0.38924	0.37367	0.36023	0.37438
Break even trans. cost	0.012%	0.000%	0.000%	0.000%	0.004%	0.000%	0.000%	0.000%	0.000%

nn = 100 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.996109	1.001800	1.011831	1.013069	1.018523	1.024860	1.033106	1.053115	1.049135
MAE	0.006075	0.006094	0.006123	0.006163	0.006189	0.006225	0.006296	0.006394	0.006384
Correl. Coeff. Act.vs.pred.	0.1117	0.0974	0.0435	0.0587	0.0480	0.0575	0.0510	0.0067	0.0466
t-value	5.94	5.18	2.31	3.12	2.55	3.06	2.71	0.36	2.48
Theil's U	0.998	0.978	0.987	0.989	1.005	0.978	1.019	1.023	1.034
Mean	0.000880	0.000716	0.000704	0.000691	0.000630	0.000633	0.000603	0.000555	0.000623
Std. Dev.	0.007467	0.007033	0.007089	0.006949	0.006469	0.006725	0.006815	0.006580	0.006246
Correct Sign Proportion	**57.08%**	55.31%	55.94%	55.10%	54.00%	54.32%	54.81%	53.72%	54.25%
Correl. Coef. sign prediction	0.0690	0.0399	0.0582	0.0404	0.0237	0.0251	0.0423	0.0173	0.0423
t-value	3.67	2.12	3.10	2.15	1.26	1.33	2.25	0.92	2.25
End wealth predict. (EW)	**1,199**	755	730	702	592	597	548	479	580
EW(pred) vs. EW(b&h)	90.18%	19.75%	15.67%	11.36%	-6.21%	-5.35%	-13.12%	-24.01%	-8.06%
Switching frequency	0.27742	0.32626	0.32767	0.37226	0.33404	0.38217	0.36306	0.32626	0.40835
Break even trans. cost	0.082%	0.020%	0.016%	0.010%	0.000%	0.000%	0.000%	0.000%	0.000%

nn = 200 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.992693	**0.989420**	0.994739	0.992388	0.994800	1.007319	1.004758	1.012565	1.011990
MAE	0.006057	**0.006055**	0.006065	0.006055	0.006074	0.006135	0.006104	0.006171	0.006159
Correl. Coeff. Act.vs.pred.	0.1256	0.1466	0.1075	0.1243	0.1136	0.0743	0.0760	0.0601	0.0634
t-value	6.67	7.80	5.71	6.61	6.04	3.95	4.04	3.19	3.37
Theil's U	1.004	1.002	0.994	0.989	1.004	0.983	0.984	0.988	0.989
Mean	0.000855	0.000845	0.000782	0.000884	0.000799	0.000632	0.000691	0.000642	0.000666
Std. Dev.	0.006592	0.006469	0.006690	0.006737	0.006454	0.006543	0.006535	0.006280	0.006420
Correct Sign Proportion	55.48%	55.59%	55.13%	56.62%	55.20%	53.64%	54.74%	53.01%	53.64%
Correl. Coef. sign prediction	0.0612	0.0626	0.0544	0.0816	0.0650	0.0428	0.0610	0.0347	0.0474
t-value	3.25	3.33	2.89	4.34	3.45	2.27	3.24	1.84	2.52
End wealth predict. (EW)	1,117	1,085	910	1,112	955	595	703	612	656
EW(pred) vs. EW(b&h)	77.13%	72.07%	44.30%	92.22%	51.38%	-5.63%	11.43%	-3.00%	4.01%
Switching frequency	0.33546	0.35103	0.31918	0.34466	0.35740	0.48832	0.46426	0.48195	0.50955
Break even trans. cost	0.060%	0.055%	0.040%	0.068%	0.041%	0.000%	0.008%	0.000%	0.003%

Non Linear predictions (Simplex method)

m =	2	3	4	5	6	7	8	9	10
NRMSE	1.161655	1.107634	1.095876	1.055095	1.040294	1.098262	1.102227	1.053511	**1.015672**
MAE	0.007290	0.007007	0.006992	0.006561	0.006439	0.007061	0.007086	0.006664	**0.006183**
Correl. Coeff. Act.vs.pred.	-0.0055	0.0454	0.0497	0.0321	0.0416	0.0301	0.0308	0.0470	**0.0593**
t-value	-0.29	2.42	2.64	1.71	2.21	1.60	1.64	2.50	3.15
Theil's U	1.198	1.088	1.062	1.074	1.061	1.117	1.173	1.092	**1.019**
Mean	0.000400	0.000521	0.000383	0.000497	0.000561	0.000238	0.000201	0.000277	0.000677
Std. Dev.	0.006262	0.005920	0.005171	0.006043	0.006250	0.004232	0.003905	0.004480	0.006827
Correct Sign Proportion	52.12%	52.72%	48.51%	51.77%	52.87%	46.18%	46.04%	49.08%	**55.52%**
Correl. Coef. sign prediction	0.0348	0.0382	-0.0089	0.0168	0.0255	-0.0173	-0.0024	0.0186	0.0604
t-value	1.85	2.03	-0.47	0.89	1.35	-0.92	-0.13	0.99	3.21
End wealth predict. (EW)	310	436	295	406	487	196	176	218	**677**
EW(pred) vs. EW(b&h)	-50.91%	-30.89%	-53.29%	-35.57%	-22.84%	-68.94%	-72.03%	-65.37%	7.33%
Switching frequency	0.51805	0.49186	0.43241	0.49894	0.47063	0.29299	0.27176	0.34607	0.38854
Break even trans. cost	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.007%

Table 8.8 : LSE Prediction Results (Library : 3- year, Prediction period : 10- year)

Linear Predictions

Prediction model	Random Walk	MA(20)	Linear model
NRMSE	1.318320	1.019351	1.004179
MAE	0.008176	0.006295	0.006126
Correl. Coeff. Act.vs.pred.	0.1310	0.0836	0.0591
t-value	6.61	4.22	2.98
Theil's U	1.000	1.069	1.002
Mean	0.000763	0.000650	0.000764
Std. Dev.	0.005915	0.005800	0.007080
Correct Sign Proportion	53.97%	52.75%	55.30%
Correl. Coef. sign prediction	0.0653	0.0239	0.0489
t-value	3.30	1.21	2.47
End wealth predict. (EW)	696	522	698
EW(pred) vs. EW(b&h)	53.37%	14.91%	53.74%
Switching frequency	0.46112	0.07306	0.32129
Break even trans. cost	0.037%	0.075%	0.053%

Market Index (Buy & hold strategy)

Mean	0.0005949
Std. Dev.	0.0087654
End Wealth	454

Non Linear predictions (Piecewise approx. method)

nn = 20 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.089239	1.076385	1.182373	1.190174	1.195420	1.245329	1.360633	1.449789	1.508037
MAE	0.006495	0.006612	0.006807	0.007001	0.007156	0.007422	0.007758	0.008233	0.008796
Correl. Coeff. Act.vs.pred.	0.0398	0.1398	0.0624	0.0832	0.0987	0.0870	0.0281	0.0017	-0.0237
t-value	2.01	7.05	3.15	4.20	4.98	4.39	1.42	0.08	-1.20
Theil's U	1.046	1.014	1.152	1.125	1.273	1.291	1.325	1.543	1.738
Mean	0.000542	0.000500	0.000505	0.000606	0.000591	0.000521	0.000404	0.000282	0.000301
Std. Dev.	0.006658	0.006991	0.007193	0.006585	0.006295	0.006505	0.006471	0.006836	0.006398
Correct Sign Proportion	54.48%	53.26%	53.61%	53.89%	53.22%	51.53%	51.57%	50.43%	50.20%
Correl. Coef. sign prediction	0.0555	0.0298	0.0374	0.0512	0.0434	0.0108	0.0189	-0.0049	-0.0048
t-value	2.80	1.51	1.89	2.58	2.19	0.55	0.95	-0.25	-0.24
End wealth predict. (EW)	462	357	361	467	449	376	279	205	215
EW(pred) vs. EW(b&h)	1.71%	-21.45%	-20.42%	2.95%	-1.07%	-17.09%	-38.52%	-54.92%	-52.58%
Switching frequency	0.46583	0.46661	0.47997	0.49018	0.48390	0.49018	0.49961	0.49568	0.47997
Break even trans. cost	0.002%	0.000%	0.000%	0.002%	0.000%	0.000%	0.000%	0.000%	0.000%

nn = 50 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.032414	1.023722	1.048224	1.048436	1.025724	1.033413	1.043856	1.064566	1.086162
MAE	0.006216	0.006217	0.006315	0.006382	0.006366	0.006357	0.006488	0.006638	0.006666
Correl. Coeff. Act.vs.pred.	-0.0078	0.1023	0.0630	0.0401	0.0965	0.1028	0.1047	0.0589	0.0714
t-value	-0.40	5.16	3.18	2.02	4.87	5.19	5.28	2.97	3.60
Theil's U	1.014	1.022	0.989	1.030	1.041	0.978	1.044	1.008	1.068
Mean	0.000595	0.000605	0.000610	0.000643	0.000689	0.000690	0.000596	0.000434	0.000533
Std. Dev.	0.007441	0.007497	0.007626	0.006908	0.006396	0.006400	0.006295	0.006384	0.005909
Correct Sign Proportion	54.83%	54.60%	54.32%	54.05%	54.60%	54.95%	53.73%	52.71%	52.59%
Correl. Coef. sign prediction	0.0430	0.0394	0.0307	0.0270	0.0455	0.0566	0.0356	0.0186	0.0205
t-value	2.17	1.99	1.55	1.36	2.30	2.86	1.80	0.94	1.03
End wealth predict. (EW)	454	466	472	514	577	578	455	302	388
EW(pred) vs. EW(b&h)	0.11%	2.56%	3.94%	13.13%	27.09%	27.41%	0.23%	-33.57%	-14.51%
Switching frequency	0.40141	0.42498	0.39592	0.39434	0.40691	0.39434	0.38649	0.42105	0.40141
Break even trans. cost	0.000%	0.002%	0.003%	0.012%	0.023%	0.024%	0.000%	0.000%	0.000%

nn = 100 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.996854	0.995824	1.004562	1.011562	1.009192	1.011947	1.018348	1.038889	1.030330
MAE	0.006144	0.006120	0.006145	0.006227	0.006220	0.006232	0.006259	0.006405	0.006385
Correl. Coeff. Act.vs.pred.	0.1101	0.1257	0.0777	0.0690	0.0865	0.0961	0.0877	0.0421	0.0855
t-value	5.55	6.34	3.92	3.48	4.37	4.85	4.42	2.13	4.31
Theil's U	1.002	0.988	0.986	0.990	0.991	0.976	0.999	1.001	0.998
Mean	0.000780	0.000732	0.000682	0.000671	0.000637	0.000649	0.000580	0.000529	0.000699
Std. Dev.	0.007477	0.007214	0.007135	0.006845	0.006602	0.006711	0.006547	0.006548	0.006267
Correct Sign Proportion	**56.87%**	56.05%	55.34%	54.79%	53.97%	54.44%	54.87%	53.65%	54.52%
Correl. Coef. sign prediction	0.0765	0.0696	0.0551	0.0492	0.0312	0.0367	0.0534	0.0285	0.0536
t-value	3.86	3.51	2.78	2.48	1.57	1.85	2.69	1.44	2.71
End wealth predict. (EW)	727	644	567	551	505	521	437	384	591
EW(pred) vs. EW(b&h)	60.04%	41.82%	24.81%	21.46%	11.33%	14.76%	-3.77%	-15.39%	30.14%
Switching frequency	0.32207	0.37863	0.38178	0.40691	0.37156	0.38570	0.39670	0.35742	0.41163
Break even trans. cost	0.058%	0.036%	0.023%	0.018%	0.011%	0.014%	0.000%	0.000%	0.025%

nn = 200 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.992531	**0.988769**	0.994396	0.992023	0.991931	1.003940	1.000229	1.010323	1.004484
MAE	0.006120	**0.006119**	0.006136	0.006125	0.006132	0.006178	0.006158	0.006229	0.006188
Correl. Coeff. Act.vs.pred.	0.1241	**0.1505**	0.1097	0.1267	0.1301	0.0844	0.0986	0.0687	0.0902
t-value	6.26	7.60	5.54	6.39	6.56	4.26	4.98	3.46	4.55
Theil's U	1.006	1.002	0.993	0.987	0.991	0.973	**0.971**	0.975	0.971
Mean	0.000792	0.000781	0.000711	0.000858	0.000813	0.000644	0.000729	0.000666	0.000793
Std. Dev.	0.006596	0.006792	0.006656	0.006491	0.006485	0.006387	0.006449	0.006217	0.006293
Correct Sign Proportion	55.22%	54.83%	53.69%	55.26%	54.71%	54.01%	55.15%	53.42%	54.91%
Correl. Coef. sign prediction	0.0624	0.0528	0.0324	0.0597	0.0544	0.0505	0.0719	0.0455	0.0734
t-value	3.15	2.66	1.64	3.01	2.75	2.55	3.63	2.30	3.70
End wealth predict. (EW)	749	728	610	**887**	790	515	639	544	**751**
EW(pred) vs. EW(b&h)	65.06%	60.39%	34.45%	95.39%	73.91%	13.44%	40.80%	19.75%	65.51%
Switching frequency	0.35035	0.36057	0.33778	0.34093	0.34643	0.45326	0.43991	0.46897	0.49018
Break even trans. cost	0.056%	0.052%	0.035%	0.078%	0.062%	0.011%	0.030%	0.015%	0.041%

Non Linear predictions (Simplex method)

m =	2	3	4	5	6	7	8	9	10
NRMSE	1.163742	1.111253	1.086190	1.043194	1.033855	1.104826	1.111283	1.060803	**1.010729**
MAE	0.007373	0.007072	0.007000	0.006527	0.006445	0.007224	0.007302	0.006823	**0.006248**
Correl. Coeff. Act.vs.pred.	-0.0304	0.0252	0.0454	0.0355	0.0428	0.0262	0.0304	0.0309	**0.0583**
t-value	-1.53	1.27	2.29	1.79	2.16	1.32	1.53	1.56	2.94
Theil's U	1.125	1.062	**0.975**	1.073	1.046	1.107	1.146	1.081	1.020
Mean	0.000311	0.000366	0.000310	0.000455	0.000454	0.000126	0.000129	0.000212	0.000558
Std. Dev.	0.006620	0.006747	0.005105	0.006122	0.006458	0.003971	0.003492	0.003819	0.007057
Correct Sign Proportion	50.98%	51.65%	48.11%	52.32%	52.24%	45.01%	44.97%	46.43%	**54.56%**
Correl. Coef. sign prediction	0.0029	0.0147	-0.0168	0.0242	0.0134	-0.0402	-0.0217	0.0051	0.0347
t-value	0.15	0.74	-0.85	1.22	0.67	-2.03	-1.10	0.26	1.75
End wealth predict. (EW)	221	254	220	318	318	138	139	172	**413**
EW(pred) vs. EW(b&h)	-51.41%	-44.07%	-51.60%	-29.89%	-30.04%	-69.65%	-69.45%	-62.22%	-8.93%
Switching frequency	0.50668	0.49254	0.41634	0.46740	0.46190	0.25609	0.20896	0.26002	0.35664
Break even trans. cost	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%

Table 8.9 : LSE Prediction Results (Library : 4- year, Prediction period : 9- year)

Linear Predictions

Prediction model	Random Walk	MA(20)	Linear model
NRMSE	1.322875	1.018548	1.003092
MAE	0.008202	0.006284	0.006125
Correl. Coeff. Act.vs.pred.	0.1250	0.0893	0.0646
t-value	5.99	4.28	3.10
Theil's U	1.000	1.069	0.997
Mean	0.000739	0.000644	0.000773
Std. Dev.	0.005912	0.005773	0.006906
Correct Sign Proportion	53.66%	52.61%	54.97%
Correl. Coef. sign prediction	0.0609	0.0243	0.0473
t-value	2.92	1.16	2.26
End wealth predict. (EW)	544	438	588
EW(pred) vs. EW(b&h)	46.60%	17.86%	58.52%
Switching frequency	0.46341	0.07404	0.33101
Break even trans. cost	0.036%	0.095%	0.061%

Market Index (Buy & hold strategy)

Mean	0.0005719
Std. Dev.	0.0088208
End Wealth	371

Non Linear predictions (Piecewise approx. method)

nn = 20 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.088513	1.063507	1.160395	1.208677	1.173136	1.364300	1.425960	1.437079	1.496865
MAE	0.006469	0.006526	0.006688	0.007008	0.007003	0.007587	0.007966	0.008218	0.008729
Correl. Coeff. Act.vs.pred.	0.0468	0.1489	0.0859	0.0013	0.1093	0.0560	-0.0025	0.0101	0.0201
t-value	2.24	7.14	4.12	0.06	5.24	2.68	-0.12	0.49	0.96
Theil's U	1.095	1.004	0.990	1.142	1.160	1.108	1.395	1.204	1.197
Mean	0.000598	0.000553	0.000423	0.000247	0.000611	0.000576	0.000415	0.000414	0.000477
Std. Dev.	0.006981	0.007206	0.007304	0.007080	0.006523	0.006364	0.006474	0.006612	0.005726
Correct Sign Proportion	54.57%	53.44%	52.13%	50.87%	53.66%	52.05%	51.39%	52.87%	51.74%
Correl. Coef. sign prediction	0.0601	0.0367	0.0079	-0.0187	0.0533	0.0178	0.0095	0.0412	0.0253
t-value	2.88	1.76	0.38	-0.90	2.55	0.85	0.45	1.98	1.21
End wealth predict. (EW)	394	356	264	176	406	374	259	259	299
EW(pred) vs. EW(b&h)	6.11%	-4.19%	-28.86%	-52.48%	9.43%	0.84%	-30.15%	-30.32%	-19.57%
Switching frequency	0.46516	0.46951	0.45732	0.45732	0.48780	0.48955	0.47300	0.47387	0.47213
Break even trans. cost	0.005%	0.000%	0.000%	0.000%	0.008%	0.000%	0.000%	0.000%	0.000%

nn = 50 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.038137	1.010839	1.046128	1.034119	1.020745	1.041851	1.065964	1.072686	1.103304
MAE	0.006220	0.006172	0.006244	0.006320	0.006361	0.006453	0.006557	0.006694	0.006787
Correl. Coeff. Act.vs.pred.	-0.0257	0.1401	0.0521	0.0672	0.1045	0.0863	0.0523	0.0404	0.0474
t-value	-1.23	6.71	2.50	3.22	5.01	4.14	2.51	1.94	2.27
Theil's U	0.997	0.981	0.998	0.987	1.084	0.981	1.019	0.964	1.059
Mean	0.000529	0.000651	0.000649	0.000526	0.000736	0.000619	0.000608	0.000490	0.000470
Std. Dev.	0.007316	0.007602	0.007578	0.007172	0.006452	0.006488	0.006206	0.006291	0.005854
Correct Sign Proportion	54.83%	54.88%	54.05%	54.66%	55.44%	53.75%	53.75%	52.53%	51.57%
Correl. Coef. sign prediction	0.0517	0.0505	0.0300	0.0527	0.0700	0.0358	0.0421	0.0221	0.0072
t-value	2.48	2.42	1.44	2.52	3.36	1.71	2.02	1.06	0.35
End wealth predict. (EW)	337	445	443	334	541	414	403	308	294
EW(pred) vs. EW(b&h)	-9.28%	19.96%	19.28%	-10.08%	45.63%	11.50%	8.69%	-17.12%	-20.86%
Switching frequency	0.41986	0.41376	0.37544	0.40854	0.40941	0.41986	0.41115	0.40679	0.42770
Break even trans. cost	0.000%	0.019%	0.020%	0.000%	0.040%	0.011%	0.009%	0.000%	0.000%

nn = 100 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.993301	0.993212	1.002801	1.008808	1.005699	1.018480	1.023968	1.043741	1.041606
MAE	0.006135	0.006106	0.006150	0.006232	0.006222	0.006278	0.006313	0.006448	0.006471
Correl. Coeff. Act.vs.pred.	0.1332	0.1384	0.0903	0.0833	0.1020	0.0843	0.0731	0.0296	0.0569
t-value	6.38	6.63	4.33	3.99	4.89	4.04	3.50	1.42	2.73
Theil's U	0.991	0.995	0.994	0.984	1.001	0.998	0.985	0.977	0.992
Mean	0.000654	0.000638	0.000681	0.000597	0.000622	0.000634	0.000580	0.000469	0.000631
Std. Dev.	0.007444	0.007140	0.007161	0.006785	0.006526	0.006511	0.006488	0.006342	0.006181
Correct Sign Proportion	55.31%	54.57%	54.66%	54.31%	53.75%	54.49%	53.75%	53.14%	53.01%
Correl. Coef. sign prediction	0.0472	0.0415	0.0440	0.0441	0.0306	0.0467	0.0343	0.0228	0.0283
t-value	2.26	1.99	2.11	2.12	1.47	2.24	1.64	1.09	1.36
End wealth predict. (EW)	448	432	477	393	416	428	378	293	425
EW(pred) vs. EW(b&h)	20.68%	16.27%	28.44%	5.86%	12.19%	15.17%	1.76%	-21.06%	14.45%
Switching frequency	0.35017	0.40157	0.37979	0.42683	0.36934	0.41812	0.41463	0.39286	0.43990
Break even trans. cost	0.023%	0.016%	0.029%	0.006%	0.013%	0.015%	0.002%	0.000%	0.013%

nn = 200 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.992267	0.991215	0.992748	0.992670	0.993791	1.009689	1.007339	1.017579	1.013534
MAE	0.006117	0.006119	0.006133	0.006133	0.006149	0.006206	0.006191	0.006268	0.006254
Correl. Coeff. Act.vs.pred.	0.1259	0.1331	0.1230	0.1236	0.1211	0.0690	0.0754	0.0493	0.0634
t-value	6.03	6.38	5.89	5.92	5.80	3.31	3.61	2.36	3.04
Theil's U	0.997	1.000	0.986	0.983	0.988	0.977	0.971	0.977	0.968
Mean	0.000739	0.000756	0.000767	0.000774	0.000774	0.000625	0.000648	0.000591	0.000713
Std. Dev.	0.006553	0.006650	0.006734	0.006577	0.006320	0.006446	0.006337	0.006361	0.006593
Correct Sign Proportion	53.96%	54.62%	54.57%	54.57%	54.44%	54.31%	54.22%	53.18%	54.27%
Correl. Coef. sign prediction	0.0461	0.0515	0.0532	0.0511	0.0560	0.0601	0.0577	0.0401	0.0605
t-value	2.21	2.47	2.55	2.45	2.68	2.88	2.77	1.92	2.90
End wealth predict. (EW)	544	566	581	590	589	420	442	388	513
EW(pred) vs. EW(b&h)	46.63%	52.46%	56.56%	58.87%	58.76%	13.04%	19.13%	4.57%	38.22%
Switching frequency	0.39286	0.35366	0.32753	0.36672	0.35714	0.45732	0.46254	0.48432	0.49129
Break even trans. cost	0.042%	0.052%	0.060%	0.055%	0.057%	0.012%	0.017%	0.004%	0.029%

Non Linear predictions (Simplex method)

m =	2	3	4	5	6	7	8	9	10
NRMSE	1.146589	1.095999	1.080245	1.046543	1.035632	1.115056	1.130643	1.072773	**1.014425**
MAE	0.007317	0.006935	0.006932	0.006542	0.006487	0.007324	0.007489	0.006930	**0.006263**
Correl. Coeff. Act.vs.pred.	-0.0209	0.0303	0.0401	0.0209	0.0228	0.0305	0.0308	0.0137	**0.0341**
t-value	-1.00	1.45	1.92	1.00	1.09	1.46	1.48	0.65	1.63
Theil's U	1.094	1.140	**1.015**	1.076	1.045	1.142	1.176	1.116	1.024
Mean	0.000248	0.000456	0.000254	0.000474	0.000426	0.000144	0.000151	0.000104	0.000503
Std. Dev.	0.006788	0.006318	0.005302	0.006220	0.006344	0.003748	0.003245	0.003546	0.007138
Correct Sign Proportion	50.17%	51.74%	47.65%	52.13%	51.09%	45.30%	45.69%	45.25%	**53.83%**
Correl. Coef. sign prediction	-0.0148	0.0173	-0.0296	0.0227	-0.0107	-0.0340	-0.0008	-0.0224	0.0226
t-value	-0.71	0.83	-1.42	1.09	-0.51	-1.63	-0.04	-1.08	1.09
End wealth predict. (EW)	177	285	179	297	265	139	142	127	**317**
EW(pred) vs. EW(b&h)	-52.38%	-23.36%	-51.72%	-20.03%	-28.50%	-62.53%	-61.88%	-65.78%	-14.57%
Switching frequency	0.49042	0.49129	0.42509	0.47648	0.46864	0.21341	0.16463	0.19338	0.35801
Break even trans. cost	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%

Table 8.10 : LSE Prediction Results (Library : 5- year, Prediction period : 8- year)

Linear Predictions

Prediction model	Random Walk	MA(20)	Linear model
NRMSE	1.321314	1.017948	1.000761
MAE	0.008399	0.006420	0.006255
Correl. Coeff. Act.vs.pred.	0.1271	0.0933	0.0753
t-value	5.75	4.22	3.41
Theil's U	1.000	1.067	0.995
Mean	0.000754	0.000651	0.000771
Std. Dev.	0.006045	0.005835	0.006989
Correct Sign Proportion	53.67%	52.64%	**55.52%**
Correl. Coef. sign prediction	0.0619	0.0274	0.0640
t-value	2.80	1.24	2.90
End wealth predict. (EW)	467	378	484
EW(pred) vs. EW(b&h)	51.54%	22.74%	56.98%
Switching frequency	0.46432	0.07136	0.33920
Break even trans. cost	0.043%	0.140%	0.065%

Market Index (Buy & hold strategy)	
Mean	0.0005506
Std. Dev.	0.0090834
End Wealth	308

Non Linear predictions (Piecewise approx. method)

nn = 20 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.079394	1.046459	1.167300	1.210449	1.171767	1.360173	1.534177	1.572925	1.658376
MAE	0.006531	0.006563	0.006869	0.007203	0.007254	0.007790	0.008368	0.008769	0.009354
Correl. Coeff. Act.vs.pred.	0.0341	**0.1659**	0.0780	0.0028	0.1053	0.0423	-0.0273	-0.0284	-0.0212
t-value	1.54	7.50	3.53	0.13	4.76	1.91	-1.23	-1.28	-0.96
Theil's U	1.068	0.980	1.180	1.093	1.277	1.218	1.490	1.406	1.384
Mean	0.000584	0.000523	0.000381	0.000516	0.000500	0.000508	0.000338	0.000305	0.000422
Std. Dev.	0.007470	0.007295	0.007455	0.006888	0.006690	0.006401	0.006665	0.006352	0.005855
Correct Sign Proportion	54.35%	52.54%	51.81%	52.98%	51.86%	51.56%	50.73%	50.93%	50.15%
Correl. Coef. sign prediction	0.0566	0.0207	0.0081	0.0381	0.0205	0.0125	0.0034	0.0106	-0.0011
t-value	2.56	0.94	0.37	1.72	0.93	0.57	0.15	0.48	-0.05
End wealth predict. (EW)	326	291	218	287	278	282	200	187	237
EW(pred) vs. EW(b&h)	5.87%	-5.58%	-29.23%	-6.82%	-9.76%	-8.38%	-35.22%	-39.45%	-23.07%
Switching frequency	0.47214	0.46334	0.48387	0.47019	0.48680	0.49853	0.51906	0.49853	0.51026
Break even trans. cost	0.006%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%

nn = 50 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.039059	1.006041	1.060793	1.025290	1.014037	1.042832	1.061573	1.083700	1.117041
MAE	0.006312	0.006268	0.006400	0.006444	0.006502	0.006606	0.006726	0.006879	0.007032
Correl. Coeff. Act.vs.pred.	-0.0280	0.1602	0.0274	0.0845	0.1245	0.0755	0.0667	0.0254	0.0495
t-value	-1.27	7.25	1.24	3.82	5.63	3.41	3.02	1.15	2.24
Theil's U	0.987	**0.967**	1.059	0.994	1.077	0.998	1.037	0.991	1.112
Mean	0.000553	0.000597	0.000625	0.000596	0.000578	0.000537	0.000556	0.000436	0.000391
Std. Dev.	0.007408	0.007798	0.007673	0.006919	0.006732	0.006369	0.006183	0.006327	0.005660
Correct Sign Proportion	54.30%	54.06%	54.30%	53.32%	53.57%	53.27%	52.20%	52.00%	50.49%
Correl. Coef. sign prediction	0.0453	0.0372	0.0449	0.0282	0.0395	0.0372	0.0233	0.0229	0.0025
t-value	2.05	1.68	2.03	1.28	1.79	1.68	1.05	1.04	0.11
End wealth predict. (EW)	309	338	359	338	326	300	311	244	222
EW(pred) vs. EW(b&h)	0.46%	9.87%	16.50%	9.79%	5.83%	-2.71%	1.10%	-20.84%	-27.88%
Switching frequency	0.45357	0.43206	0.40762	0.41544	0.43206	0.44575	0.45650	0.42326	0.44477
Break even trans. cost	ERR	0.011%	0.020%	0.011%	0.006%	0.000%	ERR	0.000%	0.000%

nn = 100 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.990458	0.990712	0.998810	1.001886	1.004175	1.014199	1.022098	1.039537	1.035544
MAE	0.006242	**0.006222**	0.006254	0.006347	0.006359	0.006416	0.006460	0.006600	0.006578
Correl. Coeff. Act.vs.pred.	0.1460	0.1459	0.1048	0.1061	0.0995	0.0899	0.0688	0.0340	0.0682
t-value	6.60	6.60	4.74	4.80	4.50	4.07	3.11	1.54	3.09
Theil's U	0.990	0.981	1.014	0.980	1.004	1.023	1.000	1.008	1.016
Mean	0.000658	0.000706	0.000603	0.000743	0.000641	0.000578	0.000537	0.000489	0.000683
Std. Dev.	0.007658	0.007276	0.007356	0.006599	0.007126	0.006641	0.006497	0.006206	0.006043
Correct Sign Proportion	54.99%	**55.33%**	54.01%	55.03%	53.86%	53.67%	52.79%	51.96%	53.03%
Correl. Coef. sign prediction	0.0434	0.0609	0.0360	0.0673	0.0389	0.0343	0.0210	0.0082	0.0423
t-value	1.96	2.75	1.63	3.04	1.76	1.55	0.95	0.37	1.91
End wealth predict. (EW)	383	423	343	**457**	371	326	300	272	403
EW(pred) vs. EW(b&h)	24.41%	37.45%	11.25%	48.23%	20.27%	5.85%	-2.71%	-11.82%	30.92%
Switching frequency	0.37732	0.39980	0.41935	0.43206	0.40665	0.43206	0.42229	0.41349	0.45455
Break even trans. cost	0.028%	0.039%	0.013%	0.045%	0.022%	0.007%	0.000%	0.000%	0.029%

nn = 200 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.991308	**0.988700**	0.990928	0.990162	0.991683	1.005340	1.005561	1.012888	1.009749
MAE	0.006239	0.006241	0.006251	0.006260	0.006285	0.006343	0.006331	0.006393	0.006361
Correl. Coeff. Act.vs.pred.	0.1337	0.1506	0.1356	0.1404	0.1330	0.0841	0.0805	0.0616	0.0740
t-value	6.05	6.81	6.13	6.35	6.02	3.80	3.64	2.78	3.35
Theil's U	0.996	0.995	0.995	0.981	0.987	0.988	0.971	0.984	0.979
Mean	0.000737	0.000796	0.000806	0.000778	0.000738	0.000668	0.000666	0.000530	0.000752
Std. Dev.	0.006344	0.006959	0.006849	0.006778	0.006386	0.006455	0.006535	0.006446	0.006391
Correct Sign Proportion	53.47%	55.08%	54.35%	54.55%	53.76%	54.20%	54.30%	52.59%	54.59%
Correl. Coef. sign prediction	0.0461	0.0693	0.0549	0.0543	0.0454	0.0614	0.0591	0.0286	0.0702
t-value	2.09	3.13	2.48	2.46	2.05	2.78	2.67	1.29	3.18
End wealth predict. (EW)	451	509	**519**	490	452	392	390	296	465
EW(pred) vs. EW(b&h)	46.25%	65.08%	68.56%	59.01%	46.70%	27.15%	26.51%	-4.05%	50.92%
Switching frequency	0.42815	0.35777	0.31867	0.36070	0.36070	0.44477	0.46041	0.48387	0.49560
Break even trans. cost	0.043%	0.068%	0.080%	0.063%	0.051%	0.026%	0.025%	0.000%	0.040%

Non Linear predictions (Simplex method)

m =	2	3	4	5	6	7	8	9	10
NRMSE	1.121545	1.083480	1.079579	1.047575	1.029138	1.120696	1.140853	1.077310	**1.011754**
MAE	0.007259	0.007007	0.007054	0.006715	0.006584	0.007519	0.007722	0.007114	**0.006416**
Correl. Coeff. Act.vs.pred.	-0.0043	0.0220	0.0304	0.0047	0.0304	0.0304	0.0165	0.0106	**0.0319**
t-value	-0.19	1.00	1.38	0.21	1.38	1.37	0.75	0.48	1.44
Theil's U	1.130	1.088	1.061	1.075	1.064	1.199	1.237	1.132	**1.001**
Mean	0.000284	0.000472	0.000216	0.000471	0.000533	0.000130	0.000128	0.000063	0.000438
Std. Dev.	0.006877	0.006570	0.005459	0.006273	0.006769	0.003436	0.002980	0.003218	0.007323
Correct Sign Proportion	50.98%	51.56%	48.29%	52.00%	**53.27%**	45.50%	45.01%	44.28%	51.86%
Correl. Coef. sign prediction	0.0030	0.0136	-0.0085	0.0248	0.0398	-0.0198	-0.0150	-0.0449	-0.0239
t-value	0.13	0.62	-0.38	1.12	1.80	-0.90	-0.68	-2.03	-1.08
End wealth predict. (EW)	179	263	155	262	**297**	130	130	114	245
EW(pred) vs. EW(b&h)	-42.03%	-14.79%	-49.53%	-14.94%	-3.50%	-57.66%	-57.82%	-63.09%	-20.56%
Switching frequency	0.50244	0.48876	0.41740	0.46921	0.48192	0.17498	0.13587	0.15836	0.36070
Break even trans. cost	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%

models, mean-variance outperform the market portfolio, however the differences found are not statistically significant.

The PW switching portfolio has the best performance, exhibiting the highest correct sign ratio ranging from 55%-58% across periods, being very close to the correct sign ratio estimated for portfolios of all the other models.

The end wealth of the best PW portfolio, exceeds the market portfolio end wealth by 59% to 120% across periods, the corresponding rates for the RW and the AR portfolios being 46%-55% and 33%-62% respectively. In addition the AR model is mean variance outperformed by the PW model.

In terms of the break even transaction cost, no model in any period exhibits a break even transaction cost which is even close to the actual transaction cost, since in all cases, the break even cost is found less than 0,1%.

As in the case of the ASE returns, the results noted above are stable and not significantly affected by the changing length of library and the prediction period.

Across prediction periods, the PW model performs better in terms of forecast measures and end wealth, for low embedding dimensions (m<5) and the highest number of neighbours (NN=200).

The SW model, compared to the other models, performs poorly, however it shows a remarkable stability since its best overall results for all periods are produced with the highest embedding parameter (m=10).

The analysis of the LSE series is concluded by testing the effect of a much longer library to the forecast results of the PW model. In the existence of chaotic dynamics, this model should be sensitive to the library length. As already mentioned, the longer data availability in the case of the LSE series gives the opportunity to perform this additional test. This has been done for two different prediction periods of 8 and 11 years with a library length of 21 and 18 years respectively. The results in comparison to those with the same prediction periods but with a much shorter library (5 and 2 years respectively) are presented in Table 8.11 below.

Table 8.11 Comparison of prediction results of the PW model (LSE series) for shorter and longer library length and fixed prediction period

Prediction period	**8**	**8**	**Change**	**11**	**11**	**Change**
Library length	**5**	**21**	**%**	**2**	**18**	**%**
NRMSE	0.9887	0.991849	**0.32%**	0.996788	0.98942	**-0.74%**
MAE	0.006222	0.006321	**1.59%**	0.006129	0.006055	**-1.21%**
CC	0.1659	0.2022	**21.88%**	0.1885	0.181	**-3.98%**
THEILS-U	0.967	0.982	**1.55%**	0.986	0.961	**-2.54%**
Correct Sign	0.5533	0.5435	**-1.77%**	0.552	0.5708	**3.41%**
End Wealth	519	509	**-1.93%**	1179	1199	**1.70%**

In Table 8.11, the improvement in 6 different indicators of the longer library forecasts over the shorter library forecasts is shown shaded. In the case of the 8-year prediction period only the CC exhibits improvement of the longer library forecasts. In the case of the 11-year prediction period most indicators are improved, but marginally. These results show clearly that nonlinear forecasts for the LSE series are not library sensitive, as expected from a non-chaotic series.

An overall visual presentation of the above findings for the ASE and the LSE series is given in Figure 8.1(a-h). In this Figure, four different indicators namely the NRMSE, the CC, the Correct Sign Proportion and the Break Even Transaction Cost are plotted for each different library length and each prediction model for the ASE (8.1a,c,e,g) and the LSE (8.1b,d,f,h) series. Through these plots, we can easily compare directly the performance among the different models, as discussed above, and especially the differences in the performance of the models between the two series. With respect to the latter, the superiority of the ASE series results is clear, indicating possible structural differences between the two markets.

Specifically, in the ASE series, it is the linear autocorrelation structure that is more prominent and determines the prediction ability of the various models employed. The chaotic component, detected through the extensive testing framework in the previous

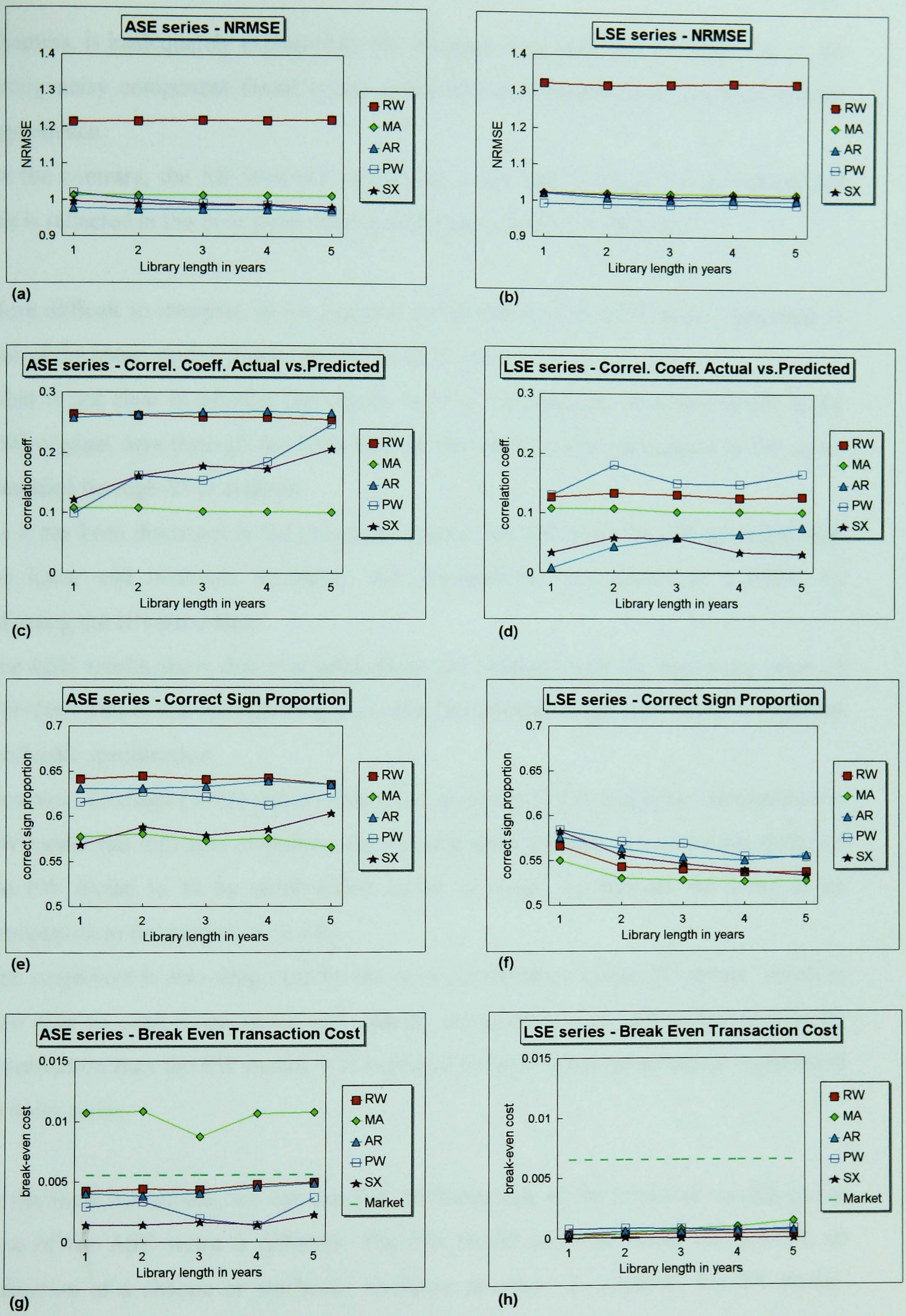

Figure 8.1 (a-h) Forecast results comparisons of each different model for alternative scenarios based on different library and predition period length. In this figure four major indicators are presented, two of them related to forecast accuracy (NRMS and Correl. coeff. between actual and predicted values) and two related to the economic assessment of forecasts (Correct sign proportion and Break-even transaction cost).

Chapters, is inadequately exploited by the nonlinear specifications, probably due to the strong noisy component found in the series and the dominance of short-term linear dependence.

On the contrary, the AR structure found also in the LSE series is less prominent and this is reflected in the poor performance of the linear models in this case.

More difficult to interpret, is the superior performance of the PW model compared to that of the linear models in the case of the LSE series.

What is not clear is whether this model exploits the nonlinear structure found in the LSE original data through the BDS test, or the strong noisy component in the series identified through SVD analysis.

As it has been discussed in the previous Chapter, the PW model is able to detect both the linear and nonlinear stochastic and deterministic components in a series, by adjusting the NN parameter.

The LSE results show that best predictions are obtained with the maximum value of NN=200. This is still inconclusive, since this NN number might also define a nonlinear stochastic specification.

However, Casdagli (1991) notices that in the presence of a strong noisy component the PW model has difficulty exploiting any existing nonlinearities. So, given the ability of the PW model to fit by construction better to noise, we support the latter as an explanation to the prediction results.

Our conjecture is also supported by the poor performance of the SX model, which is also able to exploit stochastic and chaotic components, yet, being less precise by construction than the PW model, it is expected to have worse performance when fitted to noisy series.

In the same framework, we can see that the behaviour of the nonlinear models in the case of the ASE series is different. The PW model performs better for NN=50, an indication of a chaotic or stochastic nonlinear structure. In addition, the SX model, performs better than the PW model in terms of forecast measures, a possible event in the case of chaotic data, according to Sugihara and May (1990). Recall also that, as discussed in the previous Chapter, the m parameter corresponding to the best statistical measures gives an indication of the upper dimensionality bound. In the case of the ASE series this parameter was found to be m=5, while in the case of the LSE series m=10. These results are completely in line with our previous findings concerning the structural characteristics of the two series.

A final issue that should be discussed is the good performance of the RW model in the case of the ASE series. This is mostly due to the less frequent changes of sign translated into longer intervals of successive "bullish" and "bearish" periods in the Greek market, which are better predicted by a RW naïve forecast.
Indeed, during the 13-year period spanned by the ASE and the LSE data, the proportion of positive returns was 51% and 54% respectively, while the successive sign change proportion was 36% and 46%, respectively. The latter is translated to an impressive 27,5% increase in the sign change proportion between the ASE and the LSE series.

In all, our findings reveal slightly better prediction ability of the ASE returns in terms of forecasting measures. The existence of a chaotic component is indirectly supported for the ASE series but it does not offer much to forecasting. In both cases, nonlinear dependence found by the BDS test is not economically exploitable, at least by the use of the non-linear models employed in this analysis. The same conclusion holds for the linear models with the exception of the MA model for the ASE case. This rather intriguing result will especially be addressed in one of the following sections in this Chapter.

8.3.2 ASE & LSE prediction results (Sub-period analysis)

Investigating further the predictive ability in the two markets, we present here the results of the same methodology applied to specific sub-periods of our series. The intuition behind this exercise is to find out how the different models employed perform under specific situations characterised by a) high or low volatility, b) high or low sign frequency change.

LeBaron (1992) claims that low volatility periods offer improved forecasts or, in other words, during each time period examined, there exist certain "predictability pockets" characterised by low volatility. During these periods, prediction ability measured by standard statistical measures is improved over that during high volatility periods.

On the other hand, since our economic evaluation of the models' performance is based on the return sign forecasts, we introduce a second criterion, the sign frequency change. The intuition behind this second criterion, which might help in identifying "predictability pockets", is that our models may perform better in low sign change frequency conditions. Notice that, by construction, this is the case for the RW model.

"Predictability pockets" can be defined in two ways as:

a. Sub-periods during which we can get much better predictions, as measured by various forecast accuracy indicators such as NRMSE and CC. This is similar to LeBaron's approach mentioned above.
b. Sub-periods during which we can get much better economic results i.e. periods where break-even cost is higher than the average effective transaction cost of the market.

Our aim in this section is to investigate whether "predictability pockets" exist in our series in terms of forecast statistical measures and/or in economic terms and to decide whether they are related to the criteria of volatility and sign frequency change presented above. We also investigate whether increased prediction ability, in terms of forecast measures, leads to better economic results.

Methodologically, for each of our two series, we present in a first step, analytical prediction results (in the form presented in the previous section) for the highest versus the lowest volatility and sign change frequency one-year periods. The purpose is to record any significant differences that will help to investigate the existence of "predictability pockets" and to relate predictability to these criteria.

In a second step, some key indicators (NRMS error, Correlation Coefficient between actual and predicted values, Correct Sign Proportion and Break-even Cost) are produced for all sub-periods (1983-1993) for each series. Then, the Spearman Rank Correlation Coefficient[7] is used to measure the significant relationship of each model's performance (as described by the above indicators) to the two criteria of volatility and sign change frequency.

In a final step, we investigate the correlation between the two forecast accuracy measures (NRMSE and CC) with the two indicators related to economic results (Correct Sign forecast and Break-even cost).

In order to select the highest and lowest volatility periods, our series were divided into 13 one (calendar) year periods for which annualised volatility and sign change frequency were computed. The results are presented in Table 8.12 below.

Table 8.12 Annualised volatility and sign change frequency estimation for the ASE and the LSE series

	ASE RETURNS			**LSE RETURNS**		
YEAR	**TRADING DAYS**	**VOLAT.**	**SIGN CHANGE FREQ.**	**TRADING DAYS**	**VOLAT.**	**SIGN CHANGE FREQ.**
1981	250	0.110	0.385	261	0.164	0.431
1982	250	0.100	0.378	261	0.143	0.444
1983	248	0.084	0.326	260	0.113	**0.423**
1984	250	**0.082**	0.368	261	0.131	0.444
1985	250	0.116	0.316	261	0.105	0.452
1986	248	0.151	0.298	261	0.121	0.429
1987	246	**0.577**	**0.276**	261	**0.260**	0.441
1988	251	0.184	0.367	261	0.116	0.456
1989	247	0.308	0.377	260	0.121	0.454
1990	235	0.444	0.366	261	0.137	0.471
1991	248	0.269	0.383	261	0.122	0.483
1992	252	0.246	0.385	262	0.148	**0.511**
1993*	208	0.249	**0.447**	216	**0.075**	0.454

* 1/1 - 31/10/1993

[7] The Spearman rank [Hollander and Wolf (1973)] is a non-parametric method of correlation analysis which uses the ranks of the data rather than the actual data values. This method has been

Notice that for both series (in all high and low volatility cases), predictions are based on a 2-year library for comparability purposes. The choice of the 2-year library length is dictated by the lowest sign change frequency period for the LSE series in 1983, which sets a limit of two years to the maximum common (among the ASE and the LSE series) library length available.

8.3.2.1 ASE sub-period analysis

Tables 8.13 - 8.14 present our prediction results for the highest and the lowest volatility periods and Tables 8.13[8] and 8.15 prediction results for the highest and lowest sign change frequency (s.c.f.) periods of the ASE series.

In the case of the **highest volatility** and **lowest sign change frequency period**, Table 8.13 shows clear indications of high predictability by all models. Nonlinear models perform, in general, better than the linear ones. The best NRMSE of the PW and the SX models are 3%-22% better than the corresponding estimates of the linear models. In terms of MAE the AR model is not outperformed by the nonlinear models, yet the RW and the MA model are outperformed by 10%-15%. The CC for the PW model is 7% to 500% better than the linear models but the SX model outperforms only the MA model. Finally, nonlinear models also exhibit better Theil's U statistic.

The PW model is superior in economic terms, as well. It predicts the sign of the ASE returns correctly 74 times out of 100, a very significant result, which translates into a 318% excess end wealth over the market portfolio. More important is the fact that the PW portfolio exhibits a 26% better end wealth than that of the RW model, which is favoured in periods with a very low switching frequency. Indeed, as Table 9.12 shows, the switching frequency for the ASE returns during 1987 was just 0.27, implying that this year was characterised by successive but infrequent ups and downs which enhance the predictive ability of the RW model. However, it seems that the PW model offers much better market timing opportunities, exploiting probably the chaotic component in the series.

---

selected due to the diversity of the indicators used and to the extreme values that some of them assume.

[8] Recall that the highest volatility period coincides with the lowest s.c.f. period for the ASE series.

On the other hand, the SX portfolio generates lower end wealth, yet, with a lower variance too, so a direct comparison between the two nonlinear models is more complicated.

Notice that all the different switching portfolios during this period outperform the market portfolio by 35% to 318% also exhibiting a lower variance.

When transaction cost is taken into account, the economic value of almost all predictions becomes apparent since most portfolios have a break even transaction cost much higher than the effective transaction cost of the market, which means that a true "predictability pocket" has been identified.

In the **lowest volatility period** (Table 8.14), nonlinear models outperform, in terms of most forecast measures, the RW and the MA models. In terms of the NRMSE these models are outperformed by their nonlinear counterparts by 8% to 18%, in terms of MAE by 9% to 28% and in terms of Theil's U statistic by 2% to 12%. In terms of CC only the MA model is outperformed by 183% to 204%, while the corresponding estimate for the RW model is almost identical to the estimates from PW and SX.

On the other hand the AR model performs similarly to the nonlinear models in terms of the forecast measures above.

In economic terms, all different models perform similarly, and the Correct Sign Proportion ranges from 57% for the MA model to 64,4% for the PW model. All models outperform the market portfolio by 8% to 22% also exhibiting a lower variance. However, end wealth is of similar magnitude for four out of five models tested. Finally, break-even transaction cost is always found to be much lower than the market average effective transaction cost. This means that in economic terms the lowest volatility period for the ASE series did not provide a "predictability pocket".

Table 8.15 presents the **highest sign change frequency period** prediction results. As forecast measures reflect, this is the lowest predictability period among the ones analysed for the ASE series. All models perform poorly, although nonlinear models exhibit slightly better overall performance. The SX model gives the lower NRMSE, however its difference from the rest of the models does not exceed 1%.

Table 8.13 : ASE Prediction Results

(Highest Volatility & Lowest Sign Change Frequency Period/Library : 2- year, Prediction period : 1- year)

Linear Predictions

Prediction model	Random Walk	MA(20)	Linear model
NRMSE	1.125839	1.018140	0.961033
MAE	0.023152	0.023481	0.020552
Correl. Coeff. Act.vs.pred.	0.3658	0.0655	0.3414
t-value	5.74	1.03	5.36
Theil's U	1.000	0.996	0.952
Mean	0.009137	0.005430	0.009244
Std. Dev.	0.029131	0.025014	0.029972
Correct Sign Proportion	72.36%	60.57%	71.14%
Correl. Coef. sign prediction	0.4306	0.1506	0.3948
t-value	6.75	2.36	6.19
End wealth predict. (EW)	920	375	944
EW(pred) vs. EW(b&h)	232.64%	35.52%	241.38%
Switching frequency	0.27642	0.04878	0.34959
Break even trans. cost	1.75%	2.50%	1.42%

Market Index (Buy & hold strategy)

Mean	0.0041782
Std. Dev.	0.0375388
End Wealth	277

Non Linear predictions (Piecewise approx. method)

nn = 20 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.228077	1.335697	1.403959	1.562273	1.720742	1.887822	1.735166	1.951344	1.958741
MAE	0.024811	0.026887	0.029914	0.032022	0.033647	0.037073	0.036248	0.038233	0.040560
Correl. Coeff. Act.vs.pred.	0.0627	0.0941	0.1552	0.0971	0.1636	0.1484	0.1847	0.1567	0.1781
t-value	0.98	1.48	2.43	1.52	2.57	2.33	2.90	2.46	2.79
Theil's U	0.985	1.021	1.027	1.057	1.176	1.234	1.205	1.147	1.316
Mean	0.004929	0.006267	0.003917	0.002939	0.004526	0.004154	0.003223	0.004422	0.003762
Std. Dev.	0.017850	0.018000	0.016570	0.016614	0.025202	0.023740	0.024251	0.024986	0.026502
Correct Sign Proportion	62.20%	65.04%	61.38%	54.47%	57.32%	57.72%	50.00%	55.28%	57.32%
Correl. Coef. sign prediction	0.2475	0.3039	0.2368	0.0950	0.1322	0.1582	0.0085	0.1115	0.1631
t-value	3.88	4.77	3.71	1.49	2.07	2.48	0.13	1.75	2.56
End wealth predict. (EW)	332	459	260	205	301	275	219	293	250
EW(pred) vs. EW(b&h)	20.02%	66.02%	-6.15%	-26.01%	8.82%	-0.58%	-20.72%	6.11%	-9.63%
Switching frequency	0.35772	0.38211	0.35772	0.36585	0.37398	0.39024	0.46341	0.45528	0.42276
Break even trans. cost	0.21%	0.54%	0.00%	0.00%	0.09%	0.00%	0.00%	0.05%	0.00%

nn = 50 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.043170	1.062607	1.093395	1.110226	1.075275	1.051396	1.044533	1.050988	1.036970
MAE	0.021736	0.021847	0.023249	0.023319	0.023738	0.023427	0.023683	0.023608	0.022718
Correl. Coeff. Act.vs.pred.	0.1791	0.2347	0.2622	0.2182	0.2699	0.2757	0.2919	0.2670	0.3025
t-value	2.81	3.68	4.11	3.42	4.23	4.32	4.58	4.19	4.74
Theil's U	0.973	0.980	0.956	1.007	1.033	0.970	0.952	0.997	0.976
Mean	0.007090	0.007463	0.005780	0.005394	0.006265	0.005824	0.006280	0.006610	0.007073
Std. Dev.	0.018588	0.018194	0.014215	0.014733	0.025341	0.028633	0.026641	0.027099	0.028186
Correct Sign Proportion	67.07%	69.11%	64.63%	65.45%	65.45%	67.07%	63.82%	66.67%	69.11%
Correl. Coef. sign prediction	0.3466	0.3811	0.2999	0.3055	0.3108	0.3336	0.2778	0.3241	0.3761
t-value	5.44	5.98	4.70	4.79	4.87	5.23	4.36	5.08	5.90
End wealth predict. (EW)	561	614	408	372	459	412	461	499	558
EW(pred) vs. EW(b&h)	102.67%	121.83%	47.54%	34.35%	65.96%	49.11%	66.56%	80.41%	101.87%
Switching frequency	0.38211	0.37398	0.34959	0.34146	0.39024	0.39837	0.40650	0.37398	0.40650
Break even trans. cost	0.75%	0.86%	0.45%	0.35%	0.53%	0.41%	0.51%	0.64%	0.70%

nn = 100 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.943618	0.933717	0.961552	0.979017	0.963983	0.984240	0.964741	0.970095	0.951670
MAE	0.020230	0.020213	0.020635	0.020800	0.020869	0.020753	0.020543	0.020714	0.020375
Correl. Coeff. Act.vs.pred.	0.3607	0.3824	0.3368	0.3255	0.3411	0.3412	0.3693	0.3527	0.3832
t-value	5.66	6.00	5.28	5.11	5.35	5.35	5.79	5.53	6.01
Theil's U	0.966	0.970	0.958	0.993	0.934	0.937	0.935	0.948	0.948
Mean	0.009913	0.010085	0.008468	0.007800	0.008628	0.008643	0.008918	0.008142	0.008787
Std. Dev.	0.028160	0.027955	0.020911	0.021430	0.029426	0.030542	0.030618	0.031038	0.030670
Correct Sign Proportion	73.58%	73.98%	71.54%	69.51%	69.11%	70.33%	70.73%	68.70%	71.54%
Correl. Coef. sign prediction	0.4549	0.4641	0.4155	0.3711	0.3601	0.3806	0.3897	0.3428	0.4080
t-value	7.14	7.28	6.52	5.82	5.65	5.97	6.11	5.38	6.40
End wealth predict. (EW)	1,110	1,157	783	666	814	816	873	723	845
EW(pred) vs. EW(b&h)	301.30%	318.32%	182.94%	140.73%	194.15%	195.20%	215.54%	161.47%	205.68%
Switching frequency	0.30081	0.32520	0.36585	0.35772	0.35772	0.36585	0.39837	0.39024	0.39024
Break even trans. cost	1.86%	1.77%	1.15%	0.99%	1.22%	1.20%	1.17%	1.00%	1.16%

nn = 200 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.935643	0.927712	0.923184	0.925032	0.937764	0.943924	0.944053	0.943065	0.936694
MAE	0.020403	0.020261	0.020410	0.020366	0.020587	0.020448	0.020458	0.020405	0.020357
Correl. Coeff. Act.vs.pred.	0.3601	0.3802	0.3926	0.3878	0.3635	0.3546	0.3538	0.3560	0.3677
t-value	5.65	5.96	6.16	6.08	5.70	5.56	5.55	5.58	5.77
Theil's U	0.971	0.958	0.967	0.969	0.906	0.940	0.920	0.926	0.911
Mean	0.009694	0.009551	0.009893	0.009781	0.008561	0.008675	0.008403	0.008366	0.008410
Std. Dev.	0.028368	0.028519	0.028026	0.028425	0.030455	0.030393	0.030633	0.030613	0.030612
Correct Sign Proportion	72.36%	72.36%	72.36%	71.54%	69.51%	70.73%	69.51%	70.33%	69.51%
Correl. Coef. sign prediction	0.4290	0.4275	0.4306	0.4093	0.3677	0.3910	0.3635	0.3794	0.3648
t-value	6.73	6.71	6.75	6.42	5.77	6.13	5.70	5.95	5.72
End wealth predict. (EW)	1,053	1,017	1,105	1,075	800	823	770	764	772
EW(pred) vs. EW(b&h)	280.57%	267.67%	299.33%	288.69%	189.39%	197.51%	178.56%	176.05%	179.00%
Switching frequency	0.31707	0.31707	0.32520	0.33333	0.35772	0.34146	0.35772	0.34959	0.37398
Break even trans. cost	1.70%	1.66%	1.72%	1.64%	1.20%	1.29%	1.16%	1.17%	1.11%

Non Linear predictions (Simplex method)

m =	2	3	4	5	6	7	8	9	10
NRMSE	0.960594	0.963827	**0.954011**	0.967281	0.971362	0.979689	0.977523	0.982687	0.981416
MAE	0.021784	0.021437	**0.021288**	0.022095	0.022355	0.022493	0.022615	0.022890	0.022665
Correl. Coeff. Act.vs.pred.	0.2969	0.2799	**0.3421**	0.3012	0.3195	0.2745	0.3076	0.2957	0.3003
t-value	4.66	4.39	5.37	4.72	5.01	4.31	4.82	4.64	4.71
Theil's U	0.988	1.188	0.974	0.955	1.036	1.051	0.961	0.964	0.891
Mean	0.007717	0.005061	0.006336	0.006100	0.008084	0.005523	0.005990	0.007018	0.006370
Std. Dev.	0.022419	0.026197	0.022654	0.021673	0.026755	0.031175	0.033522	0.031648	0.034698
Correct Sign Proportion	**66.26%**	61.79%	65.85%	63.82%	65.45%	63.82%	61.79%	60.57%	63.41%
Correl. Coef. sign prediction	0.3170	0.2042	0.2891	0.2476	0.2892	0.2516	0.1840	0.1577	0.2238
t-value	4.97	3.20	4.53	3.88	4.54	3.95	2.89	2.47	3.51
End wealth predict. (EW)	653	343	467	441	**713**	383	429	551	471
EW(pred) vs. EW(b&h)	135.92%	23.91%	68.84%	59.46%	157.84%	38.61%	55.23%	99.17%	70.22%
Switching frequency	0.39837	0.39837	0.36585	0.34959	0.38211	0.40650	0.34959	0.32520	0.12195
Break even trans. cost	0.87%	0.22%	0.58%	0.54%	1.00%	0.33%	0.51%	0.86%	1.75%

Table 8.14 : ASE Prediction Results (Lowest Volatility Period/Library : 2- year, Prediction period : 1- year)

Linear Predictions

Prediction model	Random Walk	MA(20)	Linear model
NRMSE	1.067749	1.012984	0.918460
MAE	0.004067	0.003655	0.003338
Correl. Coeff. Act.vs.pred.	0.4310	0.1413	0.4110
t-value	6.81	2.23	6.50
Theil's U	1.000	1.032	0.993
Mean	0.001128	0.000663	0.001025
Std. Dev.	0.003623	0.004116	0.003482
Correct Sign Proportion	62.80%	56.80%	61.60%
Correl. Coef. sign prediction	0.2526	0.1257	0.2476
t-value	3.99	1.99	3.92
End wealth predict. (EW)	**132**	118	129
EW(pred) vs. EW(b&h)	21.36%	8.14%	18.30%
Switching frequency	0.36800	0.10400	0.40000
Break even trans. cost	0.21%	0.30%	0.17%

Market Index (Buy & hold strategy)	
Mean	0.000347
Std. Dev.	0.0052205
End Wealth	109

Non Linear predictions (Piecewise approx. method)

nn = 20 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.002207	1.006134	1.122327	1.126297	1.131241	1.153764	1.182916	1.272219	1.414145
MAE	0.003731	0.003585	0.003869	0.003992	0.004048	0.004140	0.004462	0.004692	0.005325
Correl. Coeff. Act.vs.pred.	0.3061	0.3296	0.2903	0.2960	0.3059	0.2619	0.3070	0.2168	0.1696
t-value	4.84	5.21	4.59	4.68	4.84	4.14	4.85	3.43	2.68
Theil's U	1.037	1.033	1.021	1.137	**0.921**	1.084	1.043	1.148	1.016
Mean	0.000779	0.000857	0.000884	0.000720	0.000589	0.000563	0.000675	0.000721	0.000354
Std. Dev.	0.003675	0.003612	0.003527	0.003700	0.003921	0.004251	0.003857	0.003852	0.003461
Correct Sign Proportion	58.80%	64.00%	62.00%	60.00%	58.80%	57.20%	57.20%	58.80%	55.20%
Correl. Coef. sign prediction	0.1722	0.2816	0.2433	0.2119	0.1827	0.1448	0.1481	0.1759	0.1042
t-value	2.72	4.45	3.85	3.35	2.89	2.29	2.34	2.78	1.65
End wealth predict. (EW)	121	124	125	120	116	115	118	120	109
EW(pred) vs. EW(b&h)	11.29%	13.48%	14.24%	9.70%	6.19%	5.50%	8.48%	9.72%	0.18%
Switching frequency	0.40000	0.37600	0.39200	0.42400	0.43200	0.36800	0.38400	0.42400	0.43200
Break even trans. cost	0.11%	0.14%	0.14%	0.09%	0.06%	0.06%	0.08%	0.09%	0.00%

nn = 50 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.944455	0.927689	0.953350	0.992128	0.990964	0.982641	0.977488	0.977379	1.067749
MAE	0.003453	0.003366	0.003396	0.003617	0.003531	0.003574	0.003660	0.003670	0.004067
Correl. Coeff. Act.vs.pred.	0.3750	0.4120	0.3913	0.3572	0.3577	0.3680	0.3846	0.4071	0.4310
t-value	5.93	6.51	6.19	5.65	5.66	5.82	6.08	6.44	6.81
Theil's U	1.041	1.003	0.974	1.099	0.930	1.107	1.119	1.043	1.000
Mean	0.000955	0.001043	0.001022	0.000716	0.000826	0.000779	0.000754	0.000777	0.000348
Std. Dev.	0.003515	0.003346	0.003353	0.003082	0.003012	0.003692	0.003760	0.002943	0.003850
Correct Sign Proportion	60.80%	**64.40%**	64.40%	57.20%	59.20%	60.00%	56.40%	57.60%	62.80%
Correl. Coef. sign prediction	0.2154	0.3001	0.2984	0.1573	0.2021	0.2119	0.1505	0.1728	0.2526
t-value	3.41	4.75	4.72	2.49	3.20	3.35	2.38	2.73	3.99
End wealth predict. (EW)	127	130	129	119	123	121	121	121	109
EW(pred) vs. EW(b&h)	16.26%	18.83%	18.21%	9.57%	12.60%	11.30%	10.62%	11.25%	0.02%
Switching frequency	0.33600	0.35200	0.32800	0.45600	0.41600	0.42400	0.40000	0.36800	0.36800
Break even trans. cost	0.18%	0.20%	0.20%	0.08%	0.12%	0.10%	0.10%	0.12%	0.21%

nn = 100 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.927274	**0.913000**	0.919652	0.944661	0.964633	0.952599	0.959232	0.956273	0.959892
MAE	0.003365	**0.003262**	0.003285	0.003422	0.003467	0.003470	0.003540	0.003501	0.003519
Correl. Coeff. Act.vs.pred.	0.3995	**0.4269**	0.4176	0.3851	0.3632	0.3790	0.3845	0.3922	0.3882
t-value	6.32	6.75	6.60	6.09	5.74	5.99	6.08	6.20	6.14
Theil's U	0.997	1.000	0.989	1.076	1.075	1.158	1.207	1.204	1.220
Mean	0.000986	0.001019	0.001011	0.000799	0.000629	0.000719	0.000683	0.000607	0.000648
Std. Dev.	0.003365	0.003276	0.003257	0.003443	0.003247	0.003674	0.003654	0.003232	0.003219
Correct Sign Proportion	60.80%	64.00%	62.40%	57.60%	58.80%	56.80%	56.80%	57.60%	57.60%
Correl. Coef. sign prediction	0.2331	0.3005	0.2788	0.1745	0.2004	0.1562	0.1616	0.1824	0.1803
t-value	3.69	4.75	4.41	2.76	3.17	2.47	2.56	2.88	2.85
End wealth predict. (EW)	128	129	128	122	117	119	118	116	117
EW(pred) vs. EW(b&h)	17.16%	18.11%	17.89%	11.85%	7.24%	9.65%	8.70%	6.64%	7.73%
Switching frequency	0.37600	0.32800	0.29600	0.40000	0.42400	0.44800	0.39200	0.40800	0.43200
Break even trans. cost	0.17%	0.21%	0.22%	0.11%	0.07%	0.08%	0.09%	0.06%	0.07%

nn = 200 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.938346	0.933626	0.946337	0.985377	0.992970	0.998842	1.004955	0.992948	0.993529
MAE	0.003464	0.003426	0.003461	0.003642	0.003675	0.003697	0.003719	0.003689	0.003718
Correl. Coeff. Act.vs.pred.	0.3930	0.3997	0.3866	0.3501	0.3422	0.3381	0.3378	0.3498	0.3540
t-value	6.21	6.32	6.11	5.54	5.41	5.35	5.34	5.53	5.60
Theil's U	1.005	1.004	0.970	1.035	1.053	1.076	1.083	1.112	1.046
Mean	0.001063	0.001080	0.001019	0.000702	0.000699	0.000693	0.000711	0.000747	0.000794
Std. Dev.	0.003663	0.003632	0.003887	0.003655	0.003703	0.003706	0.003644	0.003650	0.003551
Correct Sign Proportion	62.00%	62.40%	61.20%	57.60%	58.40%	58.00%	58.00%	59.60%	59.20%
Correl. Coef. sign prediction	0.2381	0.2475	0.2197	0.1523	0.1664	0.1579	0.1598	0.1929	0.1844
t-value	3.76	3.91	3.47	2.41	2.63	2.50	2.53	3.05	2.92
End wealth predict. (EW)	130	131	129	119	119	119	119	120	122
EW(pred) vs. EW(b&h)	19.41%	19.93%	18.11%	9.19%	9.11%	8.94%	9.44%	10.42%	11.72%
Switching frequency	0.48800	0.46400	0.45600	0.55200	0.56800	0.57600	0.56800	0.54400	0.54400
Break even trans. cost	0.15%	0.16%	0.15%	0.07%	0.06%	0.06%	0.07%	0.07%	0.08%

Non Linear predictions (Simplex method)

m =	2	3	4	5	6	7	8	9	10
NRMSE	1.077542	0.989060	0.982859	**0.942782**	1.064982	1.195697	1.233628	1.321549	1.218683
MAE	0.004042	0.003593	0.003494	**0.003407**	0.004184	0.004963	0.005127	0.005654	0.005047
Correl. Coeff. Act.vs.pred.	0.2213	0.3188	0.2874	0.3742	0.3724	**0.3968**	0.3439	0.3819	0.3923
t-value	3.50	5.04	4.54	5.92	5.89	6.27	5.44	6.04	6.20
Theil's U	1.067	1.053	**0.980**	1.161	1.133	1.266	1.288	1.377	1.265
Mean	0.000685	0.000774	0.000669	0.000721	0.000468	0.000181	0.000190	0.000066	0.000115
Std. Dev.	0.003480	0.003636	0.003616	0.003295	0.002649	0.001388	0.001690	0.001046	0.001303
Correct Sign Proportion	54.00%	**60.80%**	59.60%	59.20%	51.60%	49.20%	47.60%	47.20%	47.60%
Correl. Coef. sign prediction	0.0936	0.2175	0.1867	0.2237	0.1499	0.1250	0.0423	0.0594	0.0842
t-value	1.48	3.44	2.95	3.54	2.37	1.98	0.67	0.94	1.33
End wealth predict. (EW)	119	**121**	118	120	112	105	105	102	103
EW(pred) vs. EW(b&h)	8.75%	11.16%	8.32%	9.72%	3.04%	-4.03%	-3.82%	-6.73%	-5.58%
Switching frequency	0.40800	0.42400	0.40000	0.36800	0.11200	0.05600	0.04800	0.00800	0.01600
Break even trans. cost	0.08%	0.10%	0.08%	0.10%	0.09%	0.00%	0.00%	0.00%	0.00%

Table 8.15 : ASE Prediction Results (Highest sign Change Frequency Period/Lib.: 2- year, Predict. period : 1- year)

Linear Predictions

Prediction model	Random Walk	MA(20)	Linear model
NRMSE	1.409959	1.018361	1.016490
MAE	0.016781	0.013195	0.012979
Correl. Coeff. Act.vs.pred.	0.0060	0.0378	0.0075
t-value	0.09	0.55	0.11
Theil's U	1.000	0.958	0.955
Mean	0.001046	0.001575	0.001466
Std. Dev.	0.013181	0.012420	0.012523
Correct Sign Proportion	54.81%	50.96%	56.73%
Correl. Coef. sign prediction	0.0954	0.0147	0.1394
t-value	1.38	0.21	2.01
End wealth predict. (EW)	124	138	135
EW(pred) vs. EW(b&h)	-0.64%	10.76%	8.30%
Switching frequency	0.44231	0.05769	0.39423
Break even trans. cost	0.00%	0.85%	0.10%

Market Index (Buy & hold strategy)

Mean	0.0010779
Std. Dev.	0.0173136
End Wealth	125

Non Linear predictions (Piecewise approx. method)

nn = 20 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.067600	1.082635	1.054872	1.051001	1.111236	1.335566	1.342269	1.417030	1.656657
MAE	0.013755	0.013939	0.013439	0.013593	0.014702	0.017398	0.017553	0.019296	0.021302
Correl. Coeff. Act.vs.pred.	0.0566	0.0788	0.1793	0.2268	0.1898	0.1572	0.1957	0.1743	0.1623
t-value	0.82	1.14	2.59	3.27	2.74	2.27	2.82	2.51	2.34
Theil's U	0.788	0.247	0.800	0.910	0.468	0.469	0.576	1.333	0.885
Mean	0.001050	0.000661	0.000994	0.001232	0.001920	0.001762	0.001872	0.000559	0.001294
Std. Dev.	0.012529	0.013757	0.013335	0.013697	0.011926	0.013022	0.011766	0.012009	0.011623
Correct Sign Proportion	53.85%	51.92%	56.25%	53.37%	55.29%	55.29%	55.29%	46.63%	54.33%
Correl. Coef. sign prediction	0.0812	0.0431	0.1284	0.0711	0.1070	0.1032	0.1064	-0.0640	0.0939
t-value	1.17	0.62	1.85	1.03	1.54	1.49	1.53	-0.92	1.35
End wealth predict. (EW)	124	115	123	129	148	144	147	112	131
EW(pred) vs. EW(b&h)	-0.57%	-8.22%	-1.71%	3.23%	18.92%	15.11%	17.74%	-10.13%	4.54%
Switching frequency	0.50962	0.42308	0.48077	0.45192	0.45192	0.49038	0.47115	0.50000	0.50000
Break even trans. cost	0.00%	0.00%	0.00%	0.03%	0.18%	0.14%	0.17%	0.00%	0.04%

nn = 50 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.035120	1.022809	1.024740	1.026113	1.037841	1.072096	1.086476	1.151098	1.164414
MAE	0.013186	0.012948	0.012847	0.013147	0.013413	0.014094	0.014264	0.014859	0.015357
Correl. Coeff. Act.vs.pred.	0.0694	0.0685	0.0916	0.0894	0.0805	0.0609	0.0533	0.0627	0.0908
t-value	1.00	0.99	1.32	1.29	1.16	0.88	0.77	0.90	1.31
Theil's U	1.295	0.630	0.757	0.621	0.781	0.419	0.688	1.142	0.866
Mean	0.000735	0.001359	0.001244	0.001061	0.000737	0.000314	0.000719	-0.000115	0.001163
Std. Dev.	0.015065	0.014471	0.014242	0.013831	0.013455	0.013452	0.013608	0.012679	0.011807
Correct Sign Proportion	53.85%	56.73%	58.17%	50.96%	50.96%	48.56%	51.44%	48.56%	54.33%
Correl. Coef. sign prediction	0.0714	0.1316	0.1615	0.0163	0.0173	-0.0300	0.0294	-0.0257	0.0958
t-value	1.03	1.90	2.33	0.23	0.25	-0.43	0.42	-0.37	1.38
End wealth predict. (EW)	116	132	129	124	116	107	116	98	127
EW(pred) vs. EW(b&h)	-6.82%	5.96%	3.47%	-0.34%	-6.77%	-14.55%	-7.12%	-21.77%	1.77%
Switching frequency	0.42308	0.47115	0.50000	0.46154	0.43269	0.56731	0.54808	0.54808	0.47115
Break even trans. cost	0.00%	0.06%	0.04%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%

nn = 100 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.027374	1.041239	1.040969	1.037240	1.049343	1.070545	1.056666	1.076568	1.063193
MAE	0.013155	0.013142	0.012926	0.013078	0.013197	0.013666	0.013482	0.013753	0.013627
Correl. Coeff. Act.vs.pred.	0.0422	0.0020	0.0375	0.0441	0.0311	-0.0042	0.0263	0.0225	0.0556
t-value	0.61	0.03	0.54	0.64	0.45	-0.06	0.38	0.32	0.80
Theil's U	1.009	0.593	0.642	0.611	0.598	0.608	0.824	0.867	0.862
Mean	0.000757	0.001167	0.001586	0.001294	0.001462	0.000503	0.000667	0.000368	0.001413
Std. Dev.	0.012761	0.012569	0.013597	0.013164	0.012951	0.012531	0.012947	0.013487	0.013809
Correct Sign Proportion	53.37%	55.77%	59.62%	56.73%	56.25%	51.44%	51.44%	51.44%	55.77%
Correl. Coef. sign prediction	0.0673	0.1163	0.1933	0.1350	0.1263	0.0306	0.0300	0.0278	0.1142
t-value	0.97	1.68	2.79	1.95	1.82	0.44	0.43	0.40	1.65
End wealth predict. (EW)	117	127	139	131	135	111	115	108	134
EW(pred) vs. EW(b&h)	-6.39%	1.84%	11.03%	4.55%	8.22%	-11.15%	-8.11%	-13.59%	7.14%
Switching frequency	0.54808	0.58654	0.54808	0.52885	0.49038	0.58654	0.56731	0.54808	0.54808
Break even trans. cost	0.00%	0.02%	0.09%	0.04%	0.08%	0.00%	0.00%	0.00%	0.06%

nn = 200 / m =	2	3	4	5	6	7	8	9	10
NRMSE	**1.017104**	1.027470	1.035023	1.027092	1.020788	1.036560	1.035062	1.041945	1.038523
MAE	0.013076	0.013050	0.013065	0.012978	0.012949	0.013181	0.013195	0.013281	0.013177
Correl. Coeff. Act.vs.pred.	0.0298	-0.0158	-0.0239	0.0106	0.0435	-0.0065	-0.0037	-0.0131	-0.0059
t-value	0.43	-0.23	-0.34	0.15	0.63	-0.09	-0.05	-0.19	-0.09
Theil's U	0.961	0.779	0.747	0.715	0.747	0.799	0.770	0.749	0.747
Mean	0.000889	0.001098	0.000841	0.001006	0.000813	0.000337	-0.000155	0.000395	0.000867
Std. Dev.	0.012534	0.012273	0.012860	0.012121	0.012209	0.012500	0.012120	0.012859	0.012995
Correct Sign Proportion	54.33%	55.77%	55.77%	56.73%	54.81%	51.44%	49.52%	51.92%	53.37%
Correl. Coef. sign prediction	0.0948	0.1218	0.1200	0.1394	0.1006	0.0318	-0.0047	0.0411	0.0679
t-value	1.37	1.76	1.73	2.01	1.45	0.46	-0.07	0.59	0.98
End wealth predict. (EW)	120	125	119	123	118	107	97	108	120
EW(pred) vs. EW(b&h)	-3.82%	0.41%	-4.75%	-1.47%	-5.30%	-14.15%	-22.42%	-13.11%	-4.25%
Switching frequency	0.47115	0.48077	0.51923	0.43269	0.42308	0.54808	0.52885	0.53846	0.48077
Break even trans. cost	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Non Linear predictions (Simplex method)

m =	2	3	4	5	6	7	8	9	10
NRMSE	1.151092	1.062963	1.050550	1.119695	**1.010031**	1.196090	1.028518	1.288102	1.015055
MAE	0.015001	0.013777	0.013463	0.015064	0.013141	0.016183	0.013383	0.017736	**0.012957**
Correl. Coeff. Act.vs.pred.	0.0459	0.1442	0.0639	0.1528	**0.1937**	0.0883	0.1312	0.0866	0.1260
t-value	0.66	2.08	0.92	2.20	2.79	1.27	1.89	1.25	1.82
Theil's U	1.283	0.391	0.862	0.373	0.886	**0.333**	1.147	0.383	0.803
Mean	0.001151	0.001327	0.000804	0.000640	0.001602	0.000692	0.001115	0.000003	0.001481
Std. Dev.	0.013105	0.012612	0.013765	0.007051	0.016011	0.005416	0.017012	0.000038	0.008497
Correct Sign Proportion	55.29%	**55.77%**	50.48%	50.00%	52.88%	50.48%	50.96%	49.04%	55.29%
Correl. Coef. sign prediction	0.1099	0.1192	0.0063	0.0358	0.0526	0.1100	-0.0124	0.0675	0.1568
t-value	1.58	1.72	0.09	0.52	0.76	1.59	-0.18	0.97	2.26
End wealth predict. (EW)	127	131	118	114	**139**	115	126	100	136
EW(pred) vs. EW(b&h)	1.52%	5.27%	-5.47%	-8.62%	11.38%	-7.64%	0.76%	-19.86%	8.64%
Switching frequency	0.50962	0.42308	0.50000	0.20192	0.24038	0.04808	0.13462	0.00962	0.24038
Break even trans. cost	0.02%	0.06%	0.00%	0.00%	**0.22%**	0.00%	0.03%	0.00%	0.17%

The same picture is given for MAE, which is almost identical among all models. More significant differences are observed in the CC measure, where PW and SX give estimates 5 to 37 times better than the estimates from the linear counterparts. Finally, in terms of Theil's U statistic, nonlinear models perform better, giving again estimates that do not exceed one third of the corresponding estimates of the linear models.

In terms of economic results, all models perform similarly but Correct Sign Proportion never exceeds 60%. The market portfolio is mean variance outperformed by the nonlinear models but only by 11% at best. When transaction cost is taken into account, only the MA model gives a higher break-even transaction cost than the effective one. This result is obviously due to the very low switching frequency in the predictions of this model (less than 6% compared to 35%-55% for all the other models) which imply a less active trading strategy as already mentioned.

The differences between prediction results of the lowest vs. highest volatility and the lowest s.c.f. vs. highest s.c.f. periods can give some first indications of whether these two criteria have discriminating power. This is done in Table 8.16 where prediction improvement of the lowest vs. highest periods for both criteria is measured. The selected indicators for prediction improvement measurement are NRMSE, CC, Correct Sign and Break-even cost.

The results in Table 8.16 are quite illuminating. The lowest volatility period seems to consistently give better predictions in terms of the two relative forecast measures (NRMSE and CC), although this improvement is low to marginal for NRMSE. These results are in line with LeBaron's (1992b) claim that during low volatility periods predictability is improved.

On the contrary, the highest volatility period performs better in terms of correct sign prediction and much better in terms of Break-even cost, since, in the last case, improvement of the highest volatility period over the low volatility one, exceeds 80% for all different models.

The picture is clearer in the case of the sign change frequency criterion. All indicators for all models exhibit a moderate to remarkable improvement between the lowest and the highest s.c.f. period.

Table 8.16. Prediction improvement (%) of the Lowest vs. Highest volatility period, and Lowest vs. Highest Sign Change frequency period for the ASE returns.

	INDICATOR	MODEL				
		PW	SX	RW	MA	AR
Vola Tility	NRMSE	1.08%	1.15%	5.15%	0.49%	4.17%
	CC	8.74%	16.67%	17.76%	115.27%	20.53%
	Correct sign	-12.95%	-8.24%	-13.21%	-6.22%	-13.41%
	Break Even Cost	-88.17%	-94.29%	-88.00%	-88.10%	-88.03%
Sign Change Freq.	NRMSE	9.25%	5.55%	20.15%	0.02%	5.46%
	CC	72.95%	76.61%	5,996.67%	73.28%	4,452.00%
	Correct sign	24.09%	18.81%	32.02%	18.86%	21.30%
	Break Even Cost	933.33%	695.82%	17,400%	194.12%	800.12%

This improvement is more pronounced for the CC and Correct Sign prediction measures. For the Break-even Cost this improvement is impressive exceeding 190% in all cases.

The findings given above seem to support the conjecture that "predictability pockets" exist[9] in the case of the ASE series and the low volatility or/and low sign change frequency criteria could help in identifying them. This is more pronounced for the sign change frequency criterion, which has been found to discriminate better, giving much improved predictions and economic results for the lowest over the highest s.c.f. periods.

However, this analysis is based only on specific periods results. Moreover, the highest volatility and lowest s.c.f. period occurs in 1987, the international market crash year

[9] The results in Table 8.16 imply the existence of "predictability pockets", in forecast measures terms, based mainly on the significant differences in the estimates of the CC indicator between the different periods, since for the NRMSE these differences are low to marginal. Notice, however, that the periods employed have been selected on the basis of certain criteria and do not necessarily produce the maximum performance difference for each indicator. Measuring the maximum performance difference in all periods is a more straightforward approach to identify "predictability pockets", in forecast measures terms. Following this approach, the NRMSE for the PW, SX and RW models exhibits improvement of 20%-30% and for the MA and AR models improvement of 5%-11%. Obviously, for all other indicators too, improvement is much higher than the results in Table 8.16.

that had similar repercussions for the ASE market. So, the question raised is whether the existence of "predictability pockets", especially in terms of economic results, is an isolated and "period specific" phenomenon that occurs only in market crash conditions. This question is also connected to the intriguing results of the MA model reported in the previous sections. Recall that the MA is the only model exhibiting break-even transaction cost higher than the effective one, irrespective of prediction period. As already mentioned, this is due to the very low trading frequency that this model implies. However, if such a simple model can consistently beat the ASE market, a serious matter is raised regarding the lack of a market mechanism able to neutralise this profit making ability. An alternative explanation could relate the MA performance to specific periods. By construction the MA model (like the RW model) performs best in cases of long and persisting trends of either direction. The period exhibiting the most intensive characteristics of this kind is 1987 as Figure 8.2 shows. So, there is the possibility that the longer period results (all of which include 1987), reported in the previous sections, are determined by the extraordinary economic results of the MA model during a sole period.

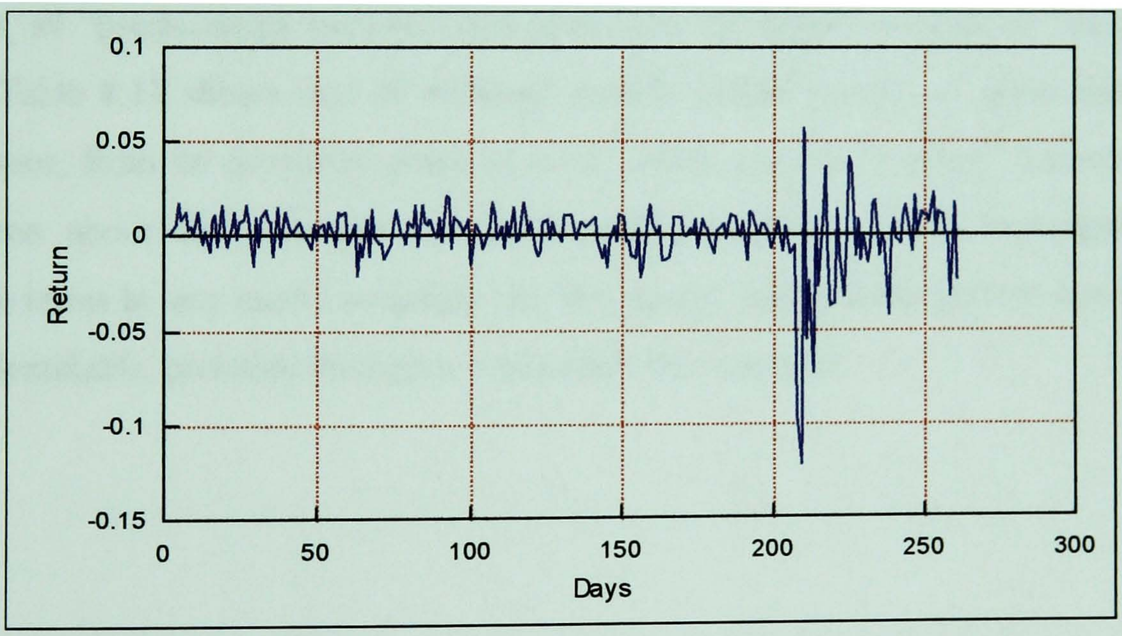

Figure 8.2. ASE : Time series plot of the 1987 annual sub-period

To clarify this issue as well as to investigate further the relationship of predictability with the criteria of volatility and sign change frequency, we analyse all the one-year sub-periods which constitute the total ASE series, as described in the second step of our methodology in section 8.3.2.

In this context, Table 8.17 presents prediction results for each model and annual sub-period. The reported indicators are related to the notion of "predictability pockets" in terms of both forecast measures and economic results.

It should be noted that the parameters used for the PW model were NN=100 and m=5, m=4 for the SX model. These parameters were chosen according to a "best prediction results" criterion based on the analytical results in Tables 8.13-8.15. Notice also that for both these models as well as for the AR(2) model used for the linear approach, a two-year library has been employed and this is why prediction results for these three models are reported after 1983.

The shaded values in Table 8.17, signify "predictability pockets" in economic terms. It is obvious that certain "predictability pockets" exist in the ASE market, during the period under study, as an outcome of the sign forecasts of different models. Most of them are attributed to the MA model for 7 different one-year periods (1982, 1983, 1987, 1990, 1991, 1992 and 1993). Hence, our previous results with respect to the MA model are not due to the isolated effect of one or two periods with high profitability.

However, all "predictability pockets" that have been identified constitute an "ex-post" finding. Table 8.17 shows that all different models exhibit periods of good and bad performance, from an economic point of view, which are not "a-priori" known. So, information about the existence of "predictability pockets" is not exploitable in economic terms by any model including the MA model, unless these periods can be "a priori" identifiable, probably through a predictable characteristic.

Table 8.17 ASE : Sub-period prediction results

Model	Indicator	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
PW	NRMSE			1.029	0.913	0.951	0.899	0.934	1.044	1.092	0.955	1.058	1.02	1.041
	CC			0.085	0.427	0.342	0.446	0.382	0.299	0.109	0.305	0.113	0.254	0.002
	Correct sign			0.58	0.64	0.64	0.69	0.74	0.59	0.63	0.66	0.6	0.56	0.56
	Break-even cost			0.35%	0.21%	0.08%	0.25%	**1.77%**	0.23%	0.18%	0.64%	0.35%	0.36%	0.02%
SX	NRMSE			0.977	0.943	1.117	0.936	0.967	1.099	1.024	1.062	1.037	1.054	1.119
	CC			0.193	0.374	0.152	0.378	0.301	0.129	0.202	0.131	0.115	0.157	0.152
	Correct sign			0.57	0.59	0.65	0.65	0.64	0.51	0.56	0.6	0.57	0.51	0.5
	Break-even cost			**0.51%**	0.10%	0.30%	0.21%	**0.54%**	0.28%	0.42%	0.21%	0.27%	0.16%	0.00%
RW	NRMSE	1.386	0.965	1.196	1.068	1.13	1.08	1.126	1.205	1.35	1.224	1.326	1.259	1.409
	CC	0.039	0.52	0.318	0.431	0.362	0.416	0.366	0.281	0.088	0.251	0.119	0.201	0.006
	Correct sign	0.61	0.62	0.67	0.63	0.68	0.7	0.72	0.63	0.62	0.63	0.62	0.61	0.55
	Break-even cost	0.18%	0.25%	**0.55%**	0.21%	0.10%	0.30%	**1.75%**	0.34%	0.12%	**0.75%**	**0.53%**	0.48%	0.00%
MA	NRMSE	1.016	1.021	1.018	1.012	1.012	1.012	1.018	1.055	1.007	1.006	1.023	1.015	1.018
	CC	0.053	0.086	0.098	0.141	0.117	0.075	0.066	-0.101	0.08	0.139	0.096	0.089	0.0378
	Correct sign	0.55	0.55	0.65	0.57	0.63	0.6	0.61	0.51	0.58	0.59	0.54	0.58	0.51
	Break-even cost	0.26%	**0.76%**	**5.60%**	0.30%	0.45%	0.00%	**2.50%**	0.15%	0.00%	**3.16%**	**1.92%**	**2.51%**	**0.85%**
AR	NRMSE			0.977	0.918	0.947	0.923	0.961	0.969	1.017	0.975	1.012	0.979	1.0164
	CC			0.284	0.411	0.345	0.389	0.341	0.301	0.102	0.224	0.105	0.203	0.0075
	Correct sign			0.64	0.62	0.69	0.7	0.71	0.59	0.61	0.63	0.54	0.62	0.57
	Break-even cost			0.44%	0.17%	0.12%	0.29%	**1.42%**	0.32%	0.11%	**0.55%**	0.28%	0.39%	0.10%

With respect to the second question, i.e. the relationship between predictability and the criteria of volatility and sign change frequency, Table 8.18 presents a Spearman Rank correlation coefficient analysis.

The hypothesis we want to test is whether periods of low volatility and s.c.f. are significantly correlated with periods exhibiting better predictability. In terms of the indicators employed and due to the descending (from the highest to the lowest value) ranking mode of the Spearman rank test, our hypothesis can be further analysed to:

- Positive autocorrelation between NRMSE and the two criteria (i.e. higher volatility or s.c.f periods correlated with higher NRMSE periods and vice-versa).
- Negative autocorrelation between the three other indicators and the two criteria (i.e. lower volatility or s.c.f periods, correlated with periods with higher CC, Correct sign and Break-even cost indicators).

Table 8.18 shows that our hypothesis is not supported as far as the volatility criterion is concerned. The RW and the AR models exhibit the expected autocorrelation sign for the first three indicators and have the highest correlation coefficients. The latter, however, do not exceed 55% and their significance level ranges from 8% to 11% being higher than the generally accepted level of 5%.

Indicators for the rest of the models exhibit low and non-significant coefficients, only one (in the case of the MA model) or two of which have the expected sign.

The picture is different for the s.c.f. criterion. All indicators for all models exhibit the expected sign and half of them exhibit significant correlation coefficients. The PW and the AR models, in particular, seem to fully support the hypothesis that low sign change frequency periods are significantly correlated with higher predictability periods. The other three models exhibit much higher correlation coefficients than in the case of the volatility criterion, yet, only for the correct sign indicator correlation is highly significant.

However, it should be noted that the higher Break-even cost periods, reflecting "predictability pockets" in economic terms, exhibit positive but not highly significant correlation with lower s.c.f. periods for all models.

Finally, an interesting remark should be made on the joint effect of the highest volatility and the lowest s.c.f. periods. The two periods that distinctively combine the two criteria

Table 8.18 ASE: Spearman rank correlation analysis between prediction indicators and the criteria of volatility and sign change frequency.

Model	Indicator	Volatility	Sign. Level	S.C.F.	Sign. Level
PW	NRMSE	0.32	0.314	0.68*	0.031
	CC	-0.17	0.58	-0.68*	0.031
	Correct Sign	0.22	0.479	-0.80*	0.011
	B-E cost	0.43	0.175	-0.32	0.306
SX	NRMSE	0.16	0.605	0.50	0.114
	CC	-0.26	0.419	-0.47	0.134
	Correct Sign	-0.10	0.750	-0.87*	0.006
	B-E cost	0.21	0.507	-0.56	0.076
RW	NRMSE	0.50	0.113	0.80*	0.011
	CC	-0.50	0.113	-0.80*	0.011
	Correct Sign	-0.16	0.610	-0.99*	0.001
	B-E cost	0.39	0.216	-0.36	0.250
MA	NRMSE	-0.05	0.883	0.22	0.482
	CC	-0.34	0.288	-0.16	0.605
	Correct Sign	-0.13	0.675	-0.78*	0.014
	B-E cost	0.18	0.574	0.03	0.976
AR	NRMSE	0.49	0.121	0.72*	0.023
	CC	-0.55	0.084	-0.75*	0.017
	Correct Sign	-0.10	0.762	-0.90*	0.004
	B-E cost	0.25	0.421	-0.48	0.128

* Statistically significant coefficient

in this opposing way are the years 1987 and 1990[10]. Notice, from table 8.17 that these two periods exhibit in almost all cases positive economic results. Of course this is only

10 As already mentioned, 1987 is the year with the highest volatility and the lowest s.c.f. The 1990 period exhibits the second highest volatility and the fifth lowest s.c.f. Hence, if we sort all the periods in a descending order according to the two criteria, 1987 and 1990 exhibit the two largest differences in their ranking.

an indication that could not be statistically supported [11]. However, intuitively, this could be true for extreme cases, combining very high volatility with very low s.c.f. since the first provide opportunities for higher profits, while for the second, sign change forecasts are much better.

A last question to be answered is whether better forecasts (in terms of NRMSE and CC) also imply better sign forecasts and/or better economic results in terms of Break-even cost. Table 8.19 shows that for three models (PW, RW and AR) forecast measures are significantly correlated with sign forecasts. On the other hand, no significant correlation is established between increased predictability, expressed by forecast measures and economic results.

Table 8.19 ASE: Spearman rank correlation analysis between forecast accuracy measures and indicators related to economic results.

Model	Indicator	NRMSE	Sign. Level	CC	Sign. level
PW	**Correct Sign**	-0.63*	0.046	0.80*	0.011
	B-E cost	-0.21	0.498	0.23	0.471
SX	**Correct Sign**	-0.45	0.152	0.39	0.220
	B-E cost	-0.21	0.517	0.06	0.851
RW	**Correct Sign**	-0.82*	0.009	0.82*	0.009
	B-E cost	-0.19	0.546	0.19	0.546
MA	**Correct Sign**	-0.36	0.257	0.37	0.236
	B-E cost	0.15	0.639	0.33	0.293
AR	**Correct Sign**	-0.63*	0.045	0.66*	0.037
	B-E cost	-0.25	0.421	0.25	0.420

* Statistically significant coefficient

11 We have used the "ranking difference" criterion to sort all the periods and then we employed the Spearman rank test to see whether there exists a positive correlation between the periods combining high volatility and low s.c.f. with the Break-even cost indicator. A positive correlation was indeed found but non-significant for all models.

However, a careful observation of the results in Table 8.17, shows that for all the sub-periods exhibiting high Break-even cost (i.e. economic results), the corresponding Correct sign estimates are also among the highest[12]. The same conclusion holds, although in a lesser degree, for the two forecast measures (NRMSE and CC) which exhibit one of their highest values during the periods with economic profitability.

So, we can argue that in the case of the ASE series a consistency exists among the different indicators used, at least with respect to the periods exhibiting economic results and higher predictability. The latter, expressed by forecast measures, is to some extent related to economic profitability.

[12] We can see in Table 8.17 that the highest Break-even cost value for the PW, RW, MA and AR models corresponds to the highest Correct sign estimate. Note that if autocorrelation analysis between the Correct sign and Break-even cost indicators is reduced to include only the periods producing economic results, autocorrelation is almost perfect. For example, autocorrelation coefficient for the MA model, which produces 6 "predictability pockets" between 1983 and 1993, exceeds 0.94.

8.3.2.2 LSE sub-period analysis

Tables 8.20 – 8.21 present prediction results for the highest and the lowest volatility periods for the LSE series. Accordingly, Tables 8.22 – 8.23 present the results for the highest and lowest s.c.f. periods.

Regarding the **highest volatility period** (Table 8.20), nonlinear models perform better than the linear ones. In terms of the NRMSE, the linear models are outperformed by the PW model, which exhibits the best prediction indicators, by 4% - 35% and by the SX model by 0.4% - 29%.

In terms of MAE, linear models are also outperformed by 4% - 28%. In terms of the CC measure, the PW model exhibits an 80% better estimate than the linear models, which however clearly outperform the SX model. Finally, in terms of the Theil's U statistic, nonlinear models are again slightly superior.

In terms of economic indicators, all models outperform mean-variance the market portfolio. The PW model performs best, exhibiting a 62% correct sign forecast that is superior by 4% to11% to the corresponding indicators for the linear models. The same model performs best in end wealth terms, exhibiting a 9% – 18% increased end wealth compared to the linear models.

After taking into account the transaction cost, only the MA model exhibits a break-even transaction cost higher than the effective market cost. As in the case of the ASE series, this is obviously due to the very low switching frequency of 6%, which limits drastically the total transaction cost of the active buy and hold simulation. In parallel, this highest volatility period coincides with the market crash year 1987, as in the case of the ASE series. Hence, as argued for the ASE case, this performance might be an isolated phenomenon, due to the exceptional market conditions during this period. This question will be addressed in the sequel.

Comparing the ASE and the LSE highest volatility periods, we observe a clear superiority of the first in terms of all indicators involved. Focusing on the relative forecast measures, the ASE period exhibits a 6% better NRMSE, a 70% better CC and a 9% better Theil's U statistic.

In terms of economic indicators, the differences are more impressive. The Correct Sign proportion is 20% better for the ASE period and in terms of the Break-even transaction cost comparison is useless since in the ASE case almost all models for all parameters produce economic results while for the LSE sub-period this happens only for the MA model. We should note, however, that the coincidence of the highest volatility with the lowest s.c.f. in the case of the ASE series may seriously affect and bias the afore mentioned comparisons in favour of the latter.

The results for the LSE **lowest volatility period** (Table 8.21) show a slight superiority of the PW nonlinear model that is trivial for NRMSE, MAE and Theil's U indicators, especially versus the AR model. However, both nonlinear models clearly outperform their linear counterparts in the case of the CC measure by more than 90%
In terms of economic results only the two nonlinear models (for specific parameters) and the RW model, outperform the market portfolio marginally. Correct Sign proportion for all models never exceeds 57% and when transaction cost is taken into account, no model produces economic results.
The ASE lowest volatility period results once again appear superior to the corresponding ones to the LSE sub-period. Comparing the best indicator values obtained among the different models, NRMSE for the ASE period is better by 9%, CC by 187% and Correct Sign proportion by 12%. Moreover, in the ASE case, the market portfolio is outperformed by 20% on average (PW, RW and AR models) when in the LSE case it is outperformed at best by 3,9% (PW model).

Prediction results for the **highest s.c.f. period** (Table 8.22) show a mixed performance picture among the various models. In terms of NRMSE the AR model outperforms the rest of the models by 1% - 31%. In terms of MAE and Theil's U statistic, differences are trivial, in terms of CC the RW model is superior by 26% - 99% and in terms of Correct Sign proportion the PW and the MA models outperform the others by 3% - 17%.
In economic terms, all models except the RW, mean-variance outperform the market portfolio by 5,6% at best and of course no model produces economic results when transaction cost is taken into account. The comparison of this period with the highest

s.c.f. ASE period shows a slight superiority of the latter in terms of most indicators. Specifically, the ASE period exhibits a superior CC indicator by 64%, a four times lower Theil's U statistic and a 3.5% better Correct Sign proportion.

However, in terms of NRMSE the LSE period appears marginally better by 2.5%. Regarding economic results, in the ASE case the market portfolio is outperformed by 11% versus 5% for the LSE sub-period.

Finally, during the **lowest s.c.f. period** (Table 8.23), PW is again the best performing model. In terms of NRMSE, it outperforms all other models by 1% - 32%, in terms of CC by 25% - 82% and in terms of Correct Sign proportion by 6% - 17%. As in the previous case, the AR is the second best performing model.

The market portfolio is mean – variance outperformed by all models except the MA. The PW model in this aspect performs the best, outperforming the market portfolio by 10%, a rather low figure. After transaction cost is accounted for, no model produces economic results and the highest Break-even cost value is lower than 0.14%.

The ASE corresponding period exhibits, in this case too, much better results. The LSE period is outperformed by 7% in terms of NRMSE, by 95% in terms of CC and by 17.5% in terms of Correct Sign proportion.

In economic terms, superiority of the ASE period is much more pronounced, however, as mentioned in the analysis of the highest volatility period, this comparison suffers from the particularity of the ASE period that is simultaneously the lowest s.c.f. and the highest volatility sub-period.

Table 8.20 : LSE Prediction Results (Highest Volatility Period/Library : 2- year, Prediction period : 1- year)

Linear Predictions

Prediction model	Random Walk	MA(20)	Linear model
NRMSE	1.310948	1.015953	1.038623
MAE	0.012434	0.009843	0.009913
Correl. Coeff. Act.vs.pred.	0.1371	0.1236	-0.0302
t-value	2.22	2.00	-0.49
Theil's U	1.000	1.221	1.020
Mean	0.000939	0.001280	0.000964
Std. Dev.	0.008556	0.007344	0.010975
Correct Sign Proportion	55.94%	57.47%	59.77%
Correl. Coef. sign prediction	0.0794	0.0770	0.0923
t-value	1.28	1.24	1.49
End wealth predict. (EW)	128	139	128
EW(pred) vs. EW(b&h)	18.10%	28.61%	18.88%
Switching frequency	0.44444	0.06130	0.28352
Break even trans. cost	0.14%	**1.85%**	0.23%

Market Index (Buy & hold strategy)	
Mean	0.0002957
Std. Dev.	0.0158288
End Wealth	108

Non Linear predictions (Piecewise approx. method)

nn = 20 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.111632	1.065021	1.288441	1.248051	1.242152	1.276417	1.248769	1.631176	1.354080
MAE	0.010171	0.010212	0.010676	0.011217	0.011751	0.012396	0.012592	0.013984	0.014807
Correl. Coeff. Act.vs.pred.	-0.1735	-0.0412	-0.1432	-0.1736	0.0948	0.0270	0.1338	0.0747	0.0371
t-value	-2.80	-0.67	-2.31	-2.81	1.53	0.44	2.16	1.21	0.60
Theil's U	1.091	1.116	1.217	1.135	1.200	1.268	1.034	1.174	**0.996**
Mean	-0.000340	-0.000065	-0.000231	-0.000600	0.000050	-0.000008	0.000710	0.000740	0.000240
Std. Dev.	0.014121	0.014323	0.014298	0.014499	0.013336	0.013145	0.008882	0.010639	0.010603
Correct Sign Proportion	50.57%	53.64%	54.79%	50.96%	51.72%	52.49%	48.66%	54.79%	49.04%
Correl. Coef. sign prediction	-0.0572	-0.0012	0.0415	-0.0366	0.0154	0.0374	-0.0250	0.0970	-0.0251
t-value	-0.92	-0.02	0.67	-0.59	0.25	0.60	-0.40	1.57	-0.41
End wealth predict. (EW)	93	98	94	86	101	100	120	121	106
EW(pred) vs. EW(b&h)	-13.89%	-8.92%	-12.75%	-20.70%	-6.18%	-7.57%	11.32%	12.19%	-1.44%
Switching frequency	0.45211	0.48276	0.49808	0.50575	0.53640	0.50575	0.50575	0.50575	0.50575
Break even trans. cost	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.08%	0.09%	0.00%

nn = 50 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.021983	1.010759	0.982497	1.002414	**0.973309**	1.024917	0.991339	0.995908	0.995666
MAE	0.009678	0.009629	0.009528	0.009714	0.009670	0.010063	0.010036	0.009891	0.010003
Correl. Coeff. Act.vs.pred.	-0.0076	0.0757	0.1973	0.1162	**0.2353**	0.0567	0.1939	0.1956	0.1919
t-value	-0.12	1.22	3.19	1.88	3.80	0.92	3.13	3.16	3.10
Theil's U	1.029	1.029	1.112	1.097	1.125	1.207	1.202	1.059	1.077
Mean	0.000650	-0.000020	-0.000330	-0.000259	0.000187	0.000282	0.001129	0.000695	0.000879
Std. Dev.	0.013235	0.014407	0.013250	0.013439	0.013221	0.011921	0.009127	0.011162	0.011025
Correct Sign Proportion	55.56%	53.26%	54.02%	53.26%	51.34%	49.81%	55.17%	54.02%	55.56%
Correl. Coef. sign prediction	0.0455	-0.0045	0.0378	0.0089	0.0060	-0.0320	0.0816	0.0688	0.0849
t-value	0.73	-0.07	0.61	0.14	0.10	-0.52	1.32	1.11	1.37
End wealth predict. (EW)	118	99	92	94	105	108	134	120	126
EW(pred) vs. EW(b&h)	9.60%	-7.85%	-14.96%	-13.39%	-2.77%	-0.36%	24.07%	10.88%	16.31%
Switching frequency	0.49042	0.50575	0.51341	0.49808	0.50575	0.51341	0.52107	0.59770	0.52107
Break even trans. cost	0.07%	0.00%	0.00%	0.00%	0.00%	0.00%	0.16%	0.07%	0.11%

nn = 100 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.013783	1.004822	0.980218	0.994814	0.984184	1.012231	0.983164	0.986565	0.985246
MAE	0.009635	0.009555	0.009460	0.009607	0.009534	0.009756	0.009703	0.009529	**0.009444**
Correl. Coeff. Act.vs.pred.	-0.0017	0.0787	0.2014	0.1267	0.1839	0.0551	0.1901	0.1847	0.1898
t-value	-0.03	1.27	3.25	2.05	2.97	0.89	3.07	2.98	3.07
Theil's U	1.030	1.032	1.078	1.065	1.090	1.099	1.051	1.096	1.061
Mean	0.000362	0.000274	0.000125	0.000467	0.000574	0.000272	0.000796	0.000614	0.000716
Std. Dev.	0.014816	0.014604	0.013595	0.013495	0.013458	0.013729	0.011126	0.011942	0.012435
Correct Sign Proportion	55.94%	55.94%	55.94%	57.47%	55.56%	53.64%	56.70%	56.70%	58.62%
Correl. Coef. sign prediction	0.0301	0.0301	0.0492	0.0741	0.0424	-0.0153	0.0690	0.0539	0.0957
t-value	0.49	0.49	0.79	1.20	0.68	-0.25	1.11	0.87	1.55
End wealth predict. (EW)	110	107	103	113	116	107	123	117	120
EW(pred) vs. EW(b&h)	1.72%	-0.57%	-4.32%	4.53%	7.48%	-0.61%	13.83%	8.59%	11.50%
Switching frequency	0.42912	0.43678	0.47510	0.42146	0.37548	0.43678	0.47510	0.42146	0.37548
Break even trans. cost	0.01%	0.00%	0.00%	0.04%	0.07%	0.00%	0.11%	0.08%	0.11%

nn = 200 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.000777	0.991953	0.979920	0.998605	0.996181	0.994174	0.985875	0.978103	0.978457
MAE	0.009689	0.009613	0.009556	0.009616	0.009659	0.009660	0.009694	0.009477	0.009501
Correl. Coeff. Act.vs.pred.	0.0578	0.1304	0.2086	0.0926	0.1101	0.1235	0.1719	0.2155	0.2161
t-value	0.93	2.11	3.37	1.50	1.78	2.00	2.78	3.48	3.49
Theil's U	1.016	1.011	1.040	1.041	1.044	1.049	1.029	1.097	1.107
Mean	0.001214	0.000569	0.000752	0.000324	0.000404	0.000823	0.000757	0.001597	0.001447
Std. Dev.	0.011211	0.012579	0.011750	0.012657	0.012788	0.011703	0.011956	0.009844	0.009841
Correct Sign Proportion	58.62%	57.09%	56.70%	56.32%	54.79%	55.94%	56.32%	**62.07%**	60.15%
Correl. Coef. sign prediction	0.1094	0.0729	0.0600	0.0405	0.0089	0.0429	0.0530	0.1881	0.1479
t-value	1.77	1.18	0.97	0.65	0.14	0.69	0.86	3.04	2.39
End wealth predict. (EW)	137	116	122	109	111	124	122	**151**	145
EW(pred) vs. EW(b&h)	26.83%	7.34%	12.55%	0.74%	2.85%	14.64%	12.69%	40.02%	34.70%
Switching frequency	0.49042	0.49808	0.44444	0.45211	0.41379	0.45211	0.45211	0.42146	0.38314
Break even trans. cost	0.18%	0.06%	0.10%	0.01%	0.03%	0.12%	0.10%	**0.31%**	0.30%

Non Linear predictions (Simplex method)

m =	2	3	4	5	6	7	8	9	10
NRMSE	1.045945	1.025286	1.016160	1.045989	1.068340	1.072042	**1.011528**	1.039347	1.258122
MAE	0.010224	0.010092	**0.009573**	0.009784	0.010229	0.010298	0.009741	0.011131	0.016156
Correl. Coeff. Act.vs.pred.	-0.0160	0.0262	**0.0551**	-0.0592	-0.0308	0.0231	-0.0259	0.0479	0.0464
t-value	-0.26	0.42	0.89	-0.96	-0.50	0.37	-0.42	0.77	0.75
Theil's U	1.166	1.295	1.186	1.185	1.300	1.265	1.105	**0.974**	1.142
Mean	-0.000424	-0.000103	0.000145	0.000210	0.000320	0.000296	0.000114	0.000000	0.000000
Std. Dev.	0.014174	0.014283	0.014959	0.015577	0.015823	0.015829	0.015427	0.000000	0.000000
Correct Sign Proportion	50.57%	53.64%	59.39%	59.77%	**60.54%**	60.15%	54.79%	40.23%	40.23%
Correl. Coef. sign prediction	-0.0361	0.0154	0.0893	0.0430	0.0762	-0.0505	-0.1051	0.0505	0.0505
t-value	-0.58	0.25	1.44	0.70	1.23	-0.82	-1.70	0.82	0.82
End wealth predict. (EW)	90	97	104	106	**109**	108	103	100	100
EW(pred) vs. EW(b&h)	-17.01%	-9.80%	-3.82%	-2.20%	0.64%	0.00%	-4.59%	-7.37%	-7.37%
Switching frequency	0.44444	0.46743	0.36015	0.14559	0.01533	0.00766	0.20690	0.00000	0.00000
Break even trans. cost	0.00%	0.00%	0.00%	0.00%	**0.14%**	0.00%	0.00%	0.00%	0.00%

Table 8.21 : LSE Prediction Results (Lowest Volatility Period/Library : 2- year, Prediction period : 1- year)

Linear Predictions

Prediction model	Random Walk	MA(20)	Linear model
NRMSE	1.360451	1.048462	0.998386
MAE	0.005420	0.004045	**0.003801**
Correl. Coeff. Act.vs.pred.	0.0756	-0.0814	0.0769
t-value	1.11	-1.20	1.13
Theil's U	1.000	1.029	0.955
Mean	0.000625	0.000223	0.000613
Std. Dev.	0.003676	0.004241	0.004830
Correct Sign Proportion	54.63%	46.76%	52.78%
Correl. Coef. sign prediction	0.0888	-0.0979	0.0184
t-value	1.30	-1.44	0.27
End wealth predict. (EW)	114	105	114
EW(pred) vs. EW(b&h)	0.23%	-8.03%	-0.02%
Switching frequency	0.46296	0.06481	0.28704
Break even trans. cost	0.00%	0.00%	0.00%

Market Index (Buy & hold strategy)	
Mean	0.000614
Std. Dev.	0.0050866
End Wealth	114

Non Linear predictions (Piecewise approx. method)

nn = 20 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.053827	1.029827	1.071813	1.087585	1.233057	1.265136	1.278071	1.306872	1.514047
MAE	0.004107	0.004036	0.004089	0.004238	0.004767	0.004905	0.004959	0.004984	0.005939
Correl. Coeff. Act.vs.pred.	0.0449	0.1410	0.1139	0.0883	-0.0739	-0.0222	0.0196	0.0416	-0.0423
t-value	0.66	2.07	1.67	1.30	-1.09	-0.33	0.29	0.61	-0.62
Theil's U	1.185	0.922	1.227	1.218	1.506	1.692	1.287	1.308	1.796
Mean	0.000557	0.000729	0.000638	0.000415	0.000255	0.000470	0.000501	0.000380	0.000205
Std. Dev.	0.004133	0.004123	0.003885	0.003970	0.003869	0.003888	0.003840	0.004101	0.003989
Correct Sign Proportion	52.31%	**57.87%**	55.56%	54.17%	50.46%	52.31%	56.02%	50.93%	49.54%
Correl. Coef. sign prediction	0.0218	0.1446	0.1015	0.0757	0.0045	0.0406	0.1172	0.0060	-0.0117
t-value	0.32	2.13	1.49	1.11	0.07	0.60	1.72	0.09	-0.17
End wealth predict. (EW)	113	117	115	109	106	111	111	108	104
EW(pred) vs. EW(b&h)	-1.20%	2.49%	0.51%	-4.17%	-7.39%	-3.03%	-2.38%	-4.89%	-8.39%
Switching frequency	0.36111	0.43519	0.47222	0.45370	0.50000	0.48148	0.47222	0.52778	0.50926
Break even trans. cost	0.00%	0.03%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

nn = 50 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.111887	1.059523	1.027922	1.037009	1.087020	1.265136	1.107730	1.085350	1.061620
MAE	0.004421	0.004155	0.004033	0.004073	0.004277	0.004905	0.004323	0.004330	0.004202
Correl. Coeff. Act.vs.pred.	-0.0108	0.0916	0.1045	0.0555	-0.0445	-0.0222	-0.0349	0.0194	0.1152
t-value	-0.16	1.35	1.54	0.82	-0.65	-0.33	-0.51	0.29	1.69
Theil's U	1.238	1.140	1.115	1.145	1.161	1.692	1.046	1.144	1.140
Mean	0.000616	0.000792	0.000656	0.000562	0.000352	0.000470	0.000223	0.000385	0.000513
Std. Dev.	0.005000	0.004274	0.004576	0.004316	0.004275	0.003888	0.003963	0.004084	0.004110
Correct Sign Proportion	53.24%	56.94%	54.17%	52.31%	50.93%	52.31%	50.93%	52.78%	55.56%
Correl. Coef. sign prediction	0.0143	0.1233	0.0589	0.0251	0.0012	0.0406	0.0108	0.0516	0.1096
t-value	0.21	1.81	0.87	0.37	0.02	0.60	0.16	0.76	1.61
End wealth predict. (EW)	114	**118**	115	113	108	111	105	109	112
EW(pred) vs. EW(b&h)	0.04%	3.87%	0.91%	-1.10%	-5.44%	-3.03%	-8.02%	-4.77%	-2.14%
Switching frequency	0.10185	0.36111	0.33333	0.32407	0.37963	0.48148	0.54630	0.49074	0.45370
Break even trans. cost	0.00%	0.05%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

nn = 100 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.012301	1.008928	1.010953	1.028345	1.039758	1.044533	1.048741	1.061074	1.041497
MAE	0.003927	0.003916	0.003914	0.004004	0.004030	0.004046	0.004091	0.004219	0.004119
Correl. Coeff. Act.vs.pred.	0.0478	0.0799	0.0735	-0.0069	-0.0257	-0.0506	-0.0281	-0.0572	0.0284
t-value	0.70	1.17	1.08	-0.10	-0.38	-0.74	-0.41	-0.84	0.42
Theil's U	0.974	1.021	1.048	1.055	0.928	1.076	0.987	1.000	0.952
Mean	0.000690	0.000754	0.000740	0.000329	0.000393	0.000110	0.000247	0.000224	0.000353
Std. Dev.	0.004949	0.004478	0.004417	0.004286	0.004263	0.004234	0.004157	0.004306	0.004034
Correct Sign Proportion	56.48%	56.48%	54.63%	48.61%	51.85%	50.00%	51.39%	47.22%	51.85%
Correl. Coef. sign prediction	0.1455	0.1125	0.0723	-0.0542	0.0135	-0.0237	0.0137	-0.0711	0.0284
t-value	2.14	1.65	1.06	-0.80	0.20	-0.35	0.20	-1.05	0.42
End wealth predict. (EW)	116	118	117	107	109	102	105	105	108
EW(pred) vs. EW(b&h)	1.63%	3.04%	2.73%	-5.92%	-4.62%	-10.23%	-7.56%	-8.00%	-5.42%
Switching frequency	0.13889	0.29630	0.40741	0.33333	0.37963	0.37963	0.38889	0.40741	0.41667
Break even trans. cost	0.05%	0.05%	0.03%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

nn = 200 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.001141	0.991793	0.996263	1.009363	1.012565	**0.996163**	1.003590	1.004930	1.005543
MAE	0.003865	0.003833	**0.003816**	0.003901	0.003912	0.003846	0.003903	0.003896	0.003884
Correl. Coeff. Act.vs.pred.	0.0896	**0.1429**	0.1065	0.0191	0.0235	0.1018	0.0636	0.0642	0.0681
t-value	1.32	2.10	1.57	0.28	0.34	1.50	0.93	0.94	1.00
Theil's U	0.950	0.979	0.962	0.963	0.987	1.018	0.942	0.941	**0.918**
Mean	0.000630	0.000620	0.000749	0.000640	0.000319	0.000641	0.000500	0.000511	0.000450
Std. Dev.	0.005064	0.005043	0.004845	0.004918	0.004653	0.004773	0.004758	0.004403	0.004365
Correct Sign Proportion	54.17%	53.24%	56.48%	54.63%	51.39%	54.17%	50.00%	53.70%	53.70%
Correl. Coef. sign prediction	0.0777	0.0158	0.1167	0.0664	-0.0084	0.0554	-0.0514	0.0544	0.0571
t-value	1.14	0.23	1.72	0.98	-0.12	0.81	-0.76	0.80	0.84
End wealth predict. (EW)	114	114	117	115	107	115	111	112	110
EW(pred) vs. EW(b&h)	0.35%	0.13%	2.92%	0.56%	-6.11%	0.58%	-2.40%	-2.17%	-3.44%
Switching frequency	0.04630	0.12037	0.28704	0.22222	0.32407	0.27778	0.30556	0.39815	0.39815
Break even trans. cost	0.00%	0.00%	**0.05%**	0.02%	0.00%	0.02%	0.00%	0.00%	0.00%

Non Linear predictions (Simplex method)

m =	2	3	4	5	6	7	8	9	10
NRMSE	1.314058	1.183717	1.138440	1.164232	1.192459	1.524444	1.592827	**1.036038**	1.813181
MAE	0.005289	0.004822	0.004506	0.004561	0.004688	0.006572	0.006655	**0.003945**	0.007980
Correl. Coeff. Act.vs.pred.	0.0447	0.1162	0.0724	0.0394	0.0356	0.0072	-0.0222	**0.1674**	0.1296
t-value	0.66	1.71	1.06	0.58	0.52	0.11	-0.33	2.46	1.90
Theil's U	1.321	1.629	1.223	1.634	1.719	1.423	2.073	**1.123**	2.273
Mean	0.000553	0.000655	0.000619	0.000347	0.000318	0.000614	0.000000	0.000444	0.000000
Std. Dev.	0.003782	0.004296	0.004202	0.002972	0.002083	0.005087	0.000000	0.002729	0.000000
Correct Sign Proportion	52.31%	51.39%	52.31%	50.00%	51.85%	52.78%	47.22%	**52.78%**	46.30%
Correl. Coef. sign prediction	0.0372	0.0196	0.0273	0.0348	**0.1360**	-0.0639	0.0639	0.1119	-0.0728
t-value	0.55	0.29	0.40	0.51	2.00	-0.94	0.94	1.64	-1.07
End wealth predict. (EW)	113	**115**	114	108	107	114	100	110	100
EW(pred) vs. EW(b&h)	-1.29%	0.88%	0.11%	-5.55%	-6.13%	0.00%	-12.31%	-3.57%	-12.31%
Switching frequency	0.48148	0.51852	0.47222	0.40741	0.17593	0.00926	0.00000	0.36111	0.00000
Break even trans. cost	0.00%	**0.01%**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Table 8.22 : LSE Prediction Results (Highest Sign Change Frequency Period/Lib.: 2- year, Pred. period : 1- year)

Linear Predictions

Prediction model	Random Walk	MA(20)	Linear model
NRMSE	1.304714	1.016653	**0.993804**
MAE	0.008932	0.006596	0.006447
Correl. Coeff. Act.vs.pred.	**0.1489**	0.0802	0.1183
t-value	2.41	1.30	1.92
Theil's U	1.000	1.193	1.009
Mean	0.000501	0.000781	0.000799
Std. Dev.	0.006975	0.006130	0.008220
Correct Sign Proportion	48.85%	57.23%	51.53%
Correl. Coef. sign prediction	-0.0244	0.1401	0.0120
t-value	-0.39	2.27	0.19
End wealth predict. (EW)	114	122	**123**
EW(pred) vs. EW(b&h)	-2.36%	5.03%	5.52%
Switching frequency	0.51908	0.04580	0.30534
Break even trans. cost	0.000%	0.400%	0.065%

Market Index (Buy & hold strategy)

Mean	0.0005927
Std. Dev.	0.0092122
End Wealth	117

Non Linear predictions (Piecewise approx. method)

nn = 20 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.103200	1.188445	1.232967	1.313439	1.363484	1.407778	1.763129	1.752454	1.786094
MAE	0.006978	0.007328	0.007504	0.008030	0.008820	0.008956	0.010528	0.010540	0.011138
Correl. Coeff. Act.vs.pred.	0.0076	-0.0019	0.0695	0.0106	0.0112	0.0705	-0.0211	0.0689	0.1122
t-value	0.12	-0.03	1.12	0.17	0.18	1.14	-0.34	1.12	1.82
Theil's U	1.207	1.090	1.184	1.382	2.240	3.428	5.406	4.101	4.636
Mean	0.000599	0.000516	0.000664	0.000421	0.000069	0.000309	-0.000158	0.000333	0.000559
Std. Dev.	0.006783	0.006692	0.006737	0.007191	0.007293	0.007287	0.007210	0.007478	0.008049
Correct Sign Proportion	52.29%	50.38%	53.05%	50.00%	47.33%	48.09%	43.51%	51.91%	50.38%
Correl. Coef. sign prediction	0.0425	0.0073	0.0611	-0.0041	-0.0569	-0.0385	-0.1343	0.0305	0.0050
t-value	0.69	0.12	0.99	-0.07	-0.92	-0.62	-2.17	0.49	0.08
End wealth predict. (EW)	117	114	119	112	102	108	96	109	116
EW(pred) vs. EW(b&h)	0.18%	-1.98%	1.86%	-4.36%	-12.72%	-7.11%	-17.72%	-6.51%	-0.86%
Switching frequency	0.49618	0.45802	0.45802	0.51145	0.55725	0.51908	0.51908	0.48855	0.53435
Break even trans. cost	0.000%	0.000%	0.016%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%

nn = 50 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.033787	1.012776	1.039471	1.061928	1.105994	1.089856	1.128144	1.153694	1.127487
MAE	0.006698	0.006606	0.006965	0.007123	0.007552	0.007465	0.007623	0.007668	0.007615
Correl. Coeff. Act.vs.pred.	0.0317	**0.1043**	0.0189	0.0174	-0.0861	-0.0264	-0.1038	-0.1416	-0.0151
t-value	0.51	1.69	0.31	0.28	-1.39	-0.43	-1.68	-2.29	-0.24
Theil's U	**1.000**	1.063	1.143	1.521	1.586	1.535	1.988	1.556	1.402
Mean	0.000468	0.000617	0.000351	0.000143	-0.000300	-0.000102	-0.000089	0.000098	0.000246
Std. Dev.	0.006462	0.005968	0.006595	0.006938	0.007091	0.006949	0.007164	0.006910	0.007055
Correct Sign Proportion	49.24%	50.00%	45.80%	46.56%	41.98%	44.27%	43.13%	46.95%	47.71%
Correl. Coef. sign prediction	-0.0079	0.0084	-0.0811	-0.0630	-0.1607	-0.1132	-0.1381	-0.0579	-0.0441
t-value	-0.13	0.14	-1.31	-1.02	-2.60	-1.83	-2.24	-0.94	-0.71
End wealth predict. (EW)	113	117	110	104	93	97	98	103	107
EW(pred) vs. EW(b&h)	-3.18%	0.64%	-6.09%	-11.03%	-20.70%	-16.51%	-16.25%	-12.06%	-8.62%
Switching frequency	0.48855	0.42748	0.51908	0.50382	0.60305	0.64885	0.60305	0.54198	0.53435
Break even trans. cost	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%

nn = 100 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.015600	1.012580	1.022824	1.043655	1.072955	1.077920	1.083611	1.101596	1.082515
MAE	**0.006398**	0.006493	0.006771	0.006909	0.007215	0.007294	0.007294	0.007455	0.007313
Correl. Coeff. Act.vs.pred.	-0.0057	0.0196	0.0402	0.0017	-0.1134	-0.0690	-0.1091	-0.0953	-0.0611
t-value	-0.09	0.32	0.65	0.03	-1.84	-1.12	-1.77	-1.54	-0.99
Theil's U	1.004	1.074	1.219	1.528	1.584	1.631	1.671	1.671	1.122
Mean	0.000723	0.000236	0.000158	0.000345	-0.000348	-0.000249	-0.000242	-0.000026	-0.000008
Std. Dev.	0.006470	0.006051	0.007128	0.007501	0.007245	0.006701	0.007109	0.007352	0.006719
Correct Sign Proportion	**57.25%**	52.29%	45.42%	49.62%	41.98%	43.89%	41.98%	43.51%	45.42%
Correl. Coef. sign prediction	0.1532	0.0464	-0.0942	-0.0085	-0.1642	-0.1228	-0.1634	-0.1343	-0.0923
t-value	2.48	0.75	-1.52	-0.14	-2.66	-1.99	-2.65	-2.17	-1.49
End wealth predict. (EW)	**121**	106	104	109	91	94	94	99	100
EW(pred) vs. EW(b&h)	3.44%	-8.86%	-10.68%	-6.22%	-21.70%	-19.65%	-19.51%	-14.86%	-14.45%
Switching frequency	0.52672	0.45038	0.35878	0.51908	0.62595	0.62595	0.65649	0.61069	0.60305
Break even trans. cost	0.025%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%

nn = 200 / m =	2	3	4	5	6	7	8	9	10
NRMSE	**1.006362**	1.009381	1.000848	1.027788	1.030636	1.025375	1.027913	1.030484	1.010477
MAE	0.006438	0.006434	0.006614	0.006733	0.006814	0.006887	0.006913	0.006944	0.006801
Correl. Coeff. Act.vs.pred.	0.0815	0.0396	0.1023	0.0064	0.0118	0.0367	0.0270	0.0385	0.1065
t-value	1.32	0.64	1.66	0.10	0.19	0.59	0.44	0.62	1.72
Theil's U	1.011	1.034	1.161	1.361	1.452	1.306	1.303	1.312	1.259
Mean	0.000670	0.000421	0.000466	-0.000158	-0.000043	0.000111	0.000114	0.000156	0.000754
Std. Dev.	0.006595	0.005865	0.006640	0.006804	0.006814	0.007572	0.007657	0.007055	0.007697
Correct Sign Proportion	52.67%	54.96%	48.47%	44.66%	45.80%	46.18%	45.04%	46.56%	51.91%
Correl. Coef. sign prediction	0.0532	0.1068	-0.0283	-0.1027	-0.0802	-0.0753	-0.0979	-0.0668	0.0379
t-value	0.86	1.73	-0.46	-1.66	-1.30	-1.22	-1.59	-1.08	0.61
End wealth predict. (EW)	**119**	112	113	96	99	103	103	104	**122**
EW(pred) vs. EW(b&h)	2.04%	-4.37%	-3.25%	-17.73%	-15.23%	-11.76%	-11.70%	-10.72%	4.28%
Switching frequency	0.29008	0.25954	0.37405	0.51145	0.54198	0.59542	0.56489	0.52672	0.48855
Break even trans. cost	0.027%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	**0.035%**

Non Linear predictions (Simplex method)

m =	2	3	4	5	6	7	8	9	10
NRMSE	1.134687	1.062510	1.073155	1.077183	1.061641	1.173289	**1.028720**	1.477799	1.123364
MAE	0.007506	0.007014	0.007052	0.007098	0.007034	0.007752	**0.006754**	0.010651	0.007474
Correl. Coeff. Act.vs.pred.	-0.0062	**0.0748**	0.0132	-0.0293	-0.0995	-0.0221	-0.0417	-0.0442	0.0449
t-value	-0.10	1.21	0.21	-0.47	-1.61	-0.36	-0.67	-0.72	0.73
Theil's U	1.505	1.469	1.260	1.518	1.609	**1.144**	1.304	2.195	1.119
Mean	0.000446	0.000657	0.000279	0.000214	0.000083	0.000007	0.000330	0.000000	0.000062
Std. Dev.	0.006975	0.006132	0.006198	0.006062	0.007400	0.001123	0.007653	0.000000	0.000976
Correct Sign Proportion	50.38%	**51.53%**	50.38%	47.71%	47.33%	48.09%	50.38%	48.09%	48.85%
Correl. Coef. sign prediction	0.0085	0.0311	0.0133	-0.0334	-0.0635	-0.0058	-0.0033	-0.0025	0.0844
t-value	0.14	0.50	0.22	-0.54	-1.03	-0.09	-0.05	-0.02	1.37
End wealth predict. (EW)	112	**119**	108	106	102	100	109	100	102
EW(pred) vs. EW(b&h)	-3.74%	1.69%	-7.82%	-9.38%	-12.42%	-14.12%	-6.59%	-14.28%	-12.89%
Switching frequency	0.52672	0.51908	0.54962	0.43511	0.51908	0.04580	0.51145	0.00000	0.01527
Break even trans. cost	0.000%	**0.012%**	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%

Table 8.23 : LSE Prediction Results (Lowest Sign Change Frequency Period/Lib.: 2-year, Pred. period : 1- year)

Linear Predictions

Prediction model	Random Walk	MA(20)	Linear model
NRMSE	1.295852	1.039408	0.991396
MAE	0.006927	0.005694	0.005428
Correl. Coeff. Act.vs.pred.	0.1602	-0.0586	0.1386
t-value	2.58	-0.95	2.24
Theil's U	1.000	0.969	1.081
Mean	0.001125	0.000717	0.001052
Std. Dev.	0.005076	0.005840	0.006682
Correct Sign Proportion	57.31%	54.23%	60.38%
Correl. Coef. sign prediction	0.1092	-0.0019	0.0708
t-value	1.76	-0.03	1.14
End wealth predict. (EW)	134	120	131
EW(pred) vs. EW(b&h)	3.75%	-6.60%	1.84%
Switching frequency	0.42308	0.04615	0.16923
Break even trans. cost	0.033%	0.000%	0.045%

Market Index (Buy & hold strategy)	
Mean	0.0009819
Std. Dev.	0.0070462
End Wealth	129

Non Linear predictions (Piecewise approx. method)

nn = 20 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.023997	1.022725	1.062737	1.039953	1.086589	1.178841	1.248036	1.309967	1.382301
MAE	0.005619	0.005698	0.005919	0.005729	0.006110	0.006529	0.006880	0.007283	0.007536
Correl. Coeff. Act.vs.pred.	0.1097	0.1234	0.0634	0.1517	0.0942	0.0116	-0.0620	-0.0125	-0.0273
t-value	1.77	1.99	1.02	2.45	1.52	0.19	-1.00	-0.20	-0.44
Theil's U	1.104	1.190	1.064	**0.941**	1.264	0.944	1.353	1.468	0.962
Mean	0.001189	0.000811	0.000809	0.000941	0.000747	0.000627	0.000235	0.000461	0.000410
Std. Dev.	0.005503	0.005747	0.005866	0.004868	0.005155	0.005364	0.005215	0.005552	0.005155
Correct Sign Proportion	60.77%	52.31%	55.38%	54.62%	50.77%	50.77%	48.46%	49.62%	48.85%
Correl. Coef. sign prediction	0.1439	-0.0104	0.0454	0.0703	0.0108	-0.0121	-0.0597	-0.0376	-0.0434
t-value	2.32	-0.17	0.73	1.13	0.17	-0.20	-0.96	-0.61	-0.70
End wealth predict. (EW)	136	123	123	127	121	118	106	113	111
EW(pred) vs. EW(b&h)	5.48%	-4.31%	-4.35%	-1.04%	-5.87%	-8.74%	-17.52%	-12.58%	-13.70%
Switching frequency	0.42308	0.53846	0.44615	0.45385	0.49231	0.54615	0.49231	0.54615	0.50000
Break even trans. cost	0.045%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%

nn = 50 / m =	2	3	4	5	6	7	8	9	10
NRMSE	1.000536	0.994793	1.019189	0.998545	1.027319	1.033475	1.090019	1.101862	1.095081
MAE	0.005479	0.005501	0.005602	0.005510	0.005713	0.005716	0.005969	0.006031	0.005999
Correl. Coeff. Act.vs.pred.	0.0879	0.1253	0.0272	0.1312	0.0299	0.0414	-0.0768	-0.0640	-0.0211
t-value	1.42	2.02	0.44	2.11	0.48	0.67	-1.24	-1.03	-0.34
Theil's U	1.052	1.143	1.109	1.153	1.165	1.247	1.417	1.377	1.175
Mean	0.001010	0.001238	0.000702	0.001200	0.000838	0.000629	0.000386	0.000497	0.000407
Std. Dev.	0.006204	0.006088	0.006122	0.006190	0.005853	0.005910	0.005949	0.005661	0.005526
Correct Sign Proportion	58.08%	61.15%	54.23%	60.38%	55.00%	52.69%	53.46%	51.15%	52.31%
Correl. Coef. sign prediction	0.0506	0.1258	-0.0207	0.1138	0.0011	-0.0526	-0.0347	-0.0366	-0.0355
t-value	0.82	2.03	-0.33	1.83	0.02	-0.85	-0.56	-0.59	-0.57
End wealth predict. (EW)	130	138	120	136	124	118	110	114	111
EW(pred) vs. EW(b&h)	0.72%	6.82%	-6.96%	5.79%	-3.64%	-8.69%	-14.25%	-11.76%	-13.77%
Switching frequency	0.39231	0.34615	0.39231	0.36154	0.38462	0.36154	0.35385	0.39231	0.35385
Break even trans. cost	0.007%	0.073%	0.000%	0.060%	0.000%	0.000%	0.000%	0.000%	0.000%

nn = 100 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.993393	1.011250	1.015135	0.997935	1.028475	1.038642	1.066353	1.063589	1.070544
MAE	0.005429	0.005626	0.005616	0.005592	0.005733	0.005754	0.005923	0.005881	0.005937
Correl. Coeff. Act.vs.pred.	0.1234	0.0577	0.0271	0.1263	0.0132	-0.0017	-0.0557	-0.0409	-0.0053
t-value	1.99	0.93	0.44	2.04	0.21	-0.03	-0.90	-0.66	-0.09
Theil's U	1.062	1.017	1.024	1.032	1.122	1.209	1.320	1.319	1.001
Mean	0.001240	0.000955	0.000932	0.001216	0.000769	0.000660	0.000570	0.000518	0.000444
Std. Dev.	0.006441	0.006012	0.006096	0.006049	0.006146	0.006239	0.006219	0.006159	0.005935
Correct Sign Proportion	60.77%	58.08%	56.54%	58.85%	53.85%	52.31%	51.54%	51.15%	51.54%
Correl. Coef. sign prediction	0.0903	0.0506	0.0230	0.0906	-0.0199	-0.0821	-0.0863	-0.0795	-0.0486
t-value	1.46	0.82	0.37	1.46	-0.32	-1.32	-1.39	-1.28	-0.78
End wealth predict. (EW)	138	128	127	137	122	119	116	114	112
EW(pred) vs. EW(b&h)	6.87%	-0.68%	-1.29%	6.21%	-5.34%	-7.96%	-10.07%	-11.27%	-12.94%
Switching frequency	0.23846	0.30000	0.33077	0.37692	0.33846	0.35385	0.31538	0.33077	0.42308
Break even trans. cost	0.107%	0.000%	0.000%	0.060%	0.000%	0.000%	0.000%	0.000%	0.000%

nn = 200 / m =	2	3	4	5	6	7	8	9	10
NRMSE	0.992235	0.988727	0.993922	**0.983708**	1.000049	1.015956	1.015208	1.018061	1.017224
MAE	0.005451	0.005424	0.005460	**0.005406**	0.005512	0.005607	0.005601	0.005624	0.005654
Correl. Coeff. Act.vs.pred.	0.1536	0.1639	0.1316	**0.1982**	0.1009	0.0565	0.0418	0.0568	0.0575
t-value	2.48	2.64	2.12	3.20	1.63	0.91	0.67	0.92	0.93
Theil's U	1.029	1.046	1.033	1.043	1.072	1.178	1.175	1.163	1.062
Mean	0.001270	0.001275	0.001011	0.001359	0.000848	0.000755	0.000764	0.000909	0.000656
Std. Dev.	0.005509	0.005598	0.005735	0.005573	0.005530	0.005918	0.005973	0.005854	0.005847
Correct Sign Proportion	61.92%	62.31%	60.00%	**63.46%**	54.62%	53.08%	54.23%	53.46%	50.38%
Correl. Coef. sign prediction	0.1643	0.1786	0.1265	0.1998	0.0386	0.0060	0.0223	0.0191	-0.0458
t-value	2.65	2.88	2.04	3.22	0.62	0.10	0.36	0.31	-0.74
End wealth predict. (EW)	139	139	130	**142**	124	121	122	126	118
EW(pred) vs. EW(b&h)	7.71%	7.85%	0.75%	10.20%	-3.38%	-5.69%	-5.46%	-1.86%	-8.05%
Switching frequency	0.28462	0.31538	0.29231	0.27692	0.35385	0.46923	0.44615	0.46154	0.51538
Break even trans. cost	0.100%	0.095%	0.009%	**0.135%**	0.000%	0.000%	0.000%	0.000%	0.000%

Non Linear predictions (Simplex method)

m =	2	3	4	5	6	7	8	9	10
NRMSE	1.175839	1.143780	1.183318	1.048998	1.016664	1.293031	1.318396	1.172736	**1.012661**
MAE	0.006695	0.006233	0.006715	0.005799	0.005674	0.007452	0.007586	0.006717	**0.005614**
Correl. Coeff. Act.vs.pred.	0.0892	0.0387	0.0541	0.0823	**0.1135**	0.0894	0.0289	0.0257	0.0522
t-value	1.44	0.62	0.87	1.33	1.83	1.44	0.47	0.41	0.84
Theil's U	0.802	1.387	0.919	1.078	1.044	**0.766**	0.805	0.891	1.042
Mean	0.001052	0.000803	0.000391	0.000582	0.001055	0.000008	0.000002	0.000003	0.000984
Std. Dev.	0.005262	0.004882	0.002986	0.004957	0.005626	0.000134	0.000120	0.000885	0.006308
Correct Sign Proportion	54.23%	56.54%	43.85%	47.69%	**55.00%**	40.00%	40.00%	39.62%	56.54%
Correl. Coef. sign prediction	0.0640	0.1115	-0.0118	-0.0587	0.0547	0.0503	0.0503	-0.0326	-0.0003
t-value	1.03	1.80	-0.19	-0.95	0.88	0.81	0.81	-0.53	-0.01
End wealth predict. (EW)	131	123	111	116	**131**	100	100	100	129
EW(pred) vs. EW(b&h)	1.82%	-4.51%	-14.13%	-9.80%	1.90%	-22.20%	-22.37%	-22.31%	0.06%
Switching frequency	0.49231	0.47692	0.33077	0.48462	0.50769	0.00769	0.00000	0.04615	0.28462
Break even trans. cost	**0.015%**	0.000%	0.000%	0.000%	0.015%	0.000%	0.000%	0.000%	0.000%

Table 8.24, as Table 8.16 in the case of the ASE sub-period analysis, presents prediction improvement of the lowest vs. the highest volatility and s.c.f. LSE periods. As we can see from this Table, in most cases the lowest volatility period seems to give worst predictions than the highest volatility period. The picture is mixed only for the AR model. Differences in favour of the highest volatility period are more pronounced for the CC and the Break-even cost indicators.

Notice that the blank cells correspond to zero or negative values of one of the indicators involved in the comparison.

Table 8.24. Prediction improvement (%) of the Lowest vs. Highest volatility period, and Lowest vs. Highest Sign Change frequency period for the LSE returns.

		MODEL				
	INDICATOR	**PW**	**SX**	**RW**	**MA**	**AR**
Vola Tility	NRMSE	-2.36%	-2.42%	-3.78%	-3.19%	3.95%
	CC	-39.3%	204.36%	-44.86%	-	-
	Correct sign	-6.77%	-12.82%	-2.34%	-18.64%	-11.70%
	Break Even Cost	-83.87%	-92.88%	-	-	-
Sign Change Freq.	NRMSE	2.23%	1.56%	0.69%	-2.16%	0.30%
	CC	90.03%	51.33%	7.52%	-	17.16%
	Correct sign	10.85%	6.73%	17.32%	-5.24%	17.17%
	Break Even Cost	285.71%	25.00%	-	-	-30.80%

The picture is quite opposite for the s.f.c. periods. As in the case of the ASE series, most models - with the exception of the MA model - seem to perform better during the lowest s.c.f. period. However, the magnitude of this difference is significantly lower than that in the case of the ASE series, as the relevant figures show.

The analysis above shows that the two criteria of volatility and s.c.f. give opposite signs. Regarding the first criterion, it is clear from Table 8.24 that, unlike the ASE case, it is the highest and not the lowest volatility period that gives better prediction results in most of the cases. Regarding the s.c.f. criterion the same Table shows that, as in the ASE case, the lowest s.c.f. period exhibits better prediction results in most of the cases. With respect to the question of whether "predictability pockets" in terms of forecast measures exist, the answer is not straightforward. The analysis reveals significant improvement of the CC indicator between the periods compared but marginal improvement of the NRMSE indicator. This picture does not change even if the improvement between the highest vs. the lowest estimate of each indicator, for all periods, is considered. In this case, the improvement in terms of the NRMSE for each model ranges between 3%-10%. Hence, "predictability pockets" can be considered only in terms of the CC measure.

From an economic point of view, the preceding analysis shows that among the four different periods that have been analysed, one can be considered as a "predictability pocket" in economic terms and for the MA model only. This period exhibits the highest volatility and includes the market crash period. So, the question of whether this finding is an isolated and period-specific phenomenon, probably due to extraordinary market conditions, is raised for the LSE case, as well.

In order to clear the picture, as we did for the ASE series, we provide prediction results for all the one-year LSE sub-periods in Table 8.25. This time the parameters used for the PW model were NN=200 and m=5, m=6 for the SX model and a AR(1) specification for the linear model. The library length and all other methodological issues are exactly the same as in the analysis presented for the ASE case.

As it becomes apparent from Table 8.25, the LSE sub-periods are less predictable than the corresponding ASE sub-periods. This is more pronounced in terms of economic results, since the only "predictability pocket" identified in the LSE case is 1987, and the only model producing these results is the MA model. Hence, the answer to the question of the previous paragraph is positive and indeed predictability in the LSE series is an isolated phenomenon most probably due to the market crash conditions of 1987 as the relevant time series plot (Figure 8.3) shows.

Table 8.25 LSE : Sub-period prediction results

Model	Indicator	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
PW	NRMSE			0.9837	0.9856	0.9987	0.9961	0.9986	0.9954	1.0033	1.0507	1.0167	1.0278	1.0094
	CC			0.1982	0.1707	0.0987	0.1274	0.093	0.1568	0.0896	-0.0855	0.0414	0.0064	0.0191
	Correct sign			0.635	0.64	0.57	0.58	0.563	0.52	0.5385	0.46	0.536	0.45	0.546
	Break-even cost			0.14%	0.16%	0.03%	0.15%	0.01%	0.02%	0.01%	-0.02%	-0.03%	-0.05%	0.02%
SX	NRMSE			1.0167	1.0324	1.1375	1.0951	1.068	1.1251	1.092	1.0594	1.1423	1.062	1.192
	CC			0.114	0.0524	0.0585	-0.0579	-0.0308	0.0654	0.053	-0.1062	0.1009	-0.0995	0.036
	Correct sign			0.55	0.582	0.49	0.533	0.606	0.471	0.585	0.506	0.498	0.473	0.519
	Break-even cost			-0.02%	-0.03%	-0.18%	-0.16%	0.14%	-0.13%	0.05%	0.06%	0.12%	-0.25%	-0.10%
RW	NRMSE	1.3314	1.3707	1.296	1.277	1.3289	1.289	1.311	1.324	1.279	1.376	1.381	1.305	1.36
	CC	0.114	0.06	0.16	0.184	0.118	0.169	0.137	0.142	0.181	0.055	0.027	0.149	0.076
	Correct sign	0.569	0.555	0.573	0.556	0.5402	0.548	0.559	0.549	0.535	0.506	0.513	0.489	0.546
	Break-even cost	0.10%	-0.01%	0.03%	0.01%	-0.01%	0.06%	0.14%	0.07%	0.00%	0.10%	-0.10%	-0.03%	0.00%
MA	NRMSE	1.0224	1.0361	1.0394	1.0273	1.0286	1.0157	1.016	1.029	1.0145	1.0284	1.012	1.0167	1.048
	CC	0.0419	-0.10912	-0.059	0.0117	0.003	0.084	0.124	0.0026	0.102	0.0214	0.0986	0.08	-0.081
	Correct sign	0.533	0.46	0.542	0.548	0.509	0.544	0.575	0.467	0.5808	0.475	0.502	0.572	0.468
	Break-even cost	-0.04%	-0.25%	-0.09%	-0.03%	-0.05%	-0.07%	**1.85%**	-0.08%	0.25%	0.47%	-0.06%	0.40%	-0.09%
AR	NRMSE			0.9914	0.987	0.9961	0.9881	1.039	0.9906	0.9862	1.0094	1.0064	0.994	0.998
	CC			0.139	0.165	0.109	0.155	-0.03	0.137	0.197	0.0317	-0.0433	0.118	0.077
	Correct sign			0.604	0.609	0.54	0.571	0.598	0.548	0.558	0.483	0.502	0.515	0.528
	Break-even cost			0.05%	0.04%	0.04%	0.10%	0.23%	0.04%	0.10%	0.06%	-0.08%	0.07%	0.00%

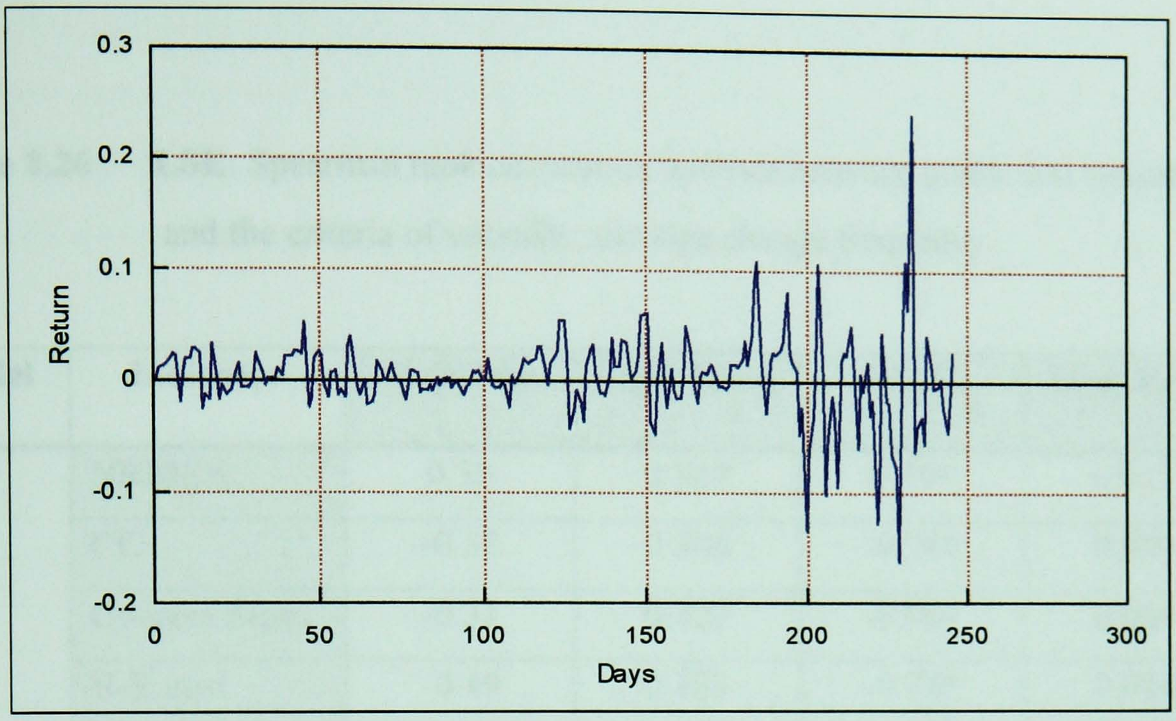

Figure 8.3. LSE: Time series plot of the 1987 annual sub-period

In Table 8.26 we provide again a Spearman-rank correlation analysis to further investigate the relationship of volatility and s.c.f criteria to increased predictability.
With respect to the volatility criterion, as in the case of the ASE series the hypothesis of correlation between low volatility periods with increased predictability is not supported. Actually, this correlation is much weaker than in the case of ASE series since the "expected" sign according to our hypothesis is correct only for the PW model but still not significant. On the contrary, the only significant correlations are for the CC and Break-even cost indicators of the MA model, both of which exhibit the opposite (to the expected) sign. Hence, the results of our previous analysis, according to which high volatility is related to increased predictability and better economic results, are verified.

On the other hand, with respect to the s.c.f. criterion, Table 8.26 shows that low s.c.f. periods are related to higher predictability. With the exception of the MA model, all other models exhibit the "correct" sign for all indicators according to our null hypothesis. In addition, for the PW model, all correlation coefficients are significant and for the SX, RW and AR models at least one coefficient is also significant. These results verify our previous findings presented in Table 8.24 and are also in line with our findings for the ASE series.

Table 8.26 LSE: Spearman rank correlation analysis between prediction indicators and the criteria of volatility and sign change frequency.

Model	Indicator	Volatility	Sign. Level	S.C.F.	Sign. Level
PW	NRMSE	0.30	0.349	0.79*	0.013
	CC	-0.32	0.306	-0.74*	0.020
	Correct Sign	-0.31	0.327	-0.90*	0.004
	B-E cost	-0.49	0.123	-0.78*	0.014
SX	NRMSE	-0.48	0.130	0.30	0.349
	CC	-0.55	0.084	-0.20	0.526
	Correct Sign	0.21	0.498	-0.65*	0.041
	B-E cost	0.37	0.243	-0.07	0.829
RW	NRMSE	-0.11	0.719	0.53	0.092
	CC	0.05	0.874	-0.49	0.123
	Correct Sign	-0.21	0.498	-0.84*	0.008
	B-E cost	0.24	0.448	-0.44	0.164
MA	NRMSE	-0.58	0.067	-0.18	0.574
	CC	0.71*	0.025	0.10	0.740
	Correct Sign	0.51	0.107	-0.28	0.379
	B-E cost	0.84*	0.008	0.27	0.398
AR	NRMSE	0.26	0.403	0.25	0.428
	CC	-0.22	0.489	-0.39	0.220
	Correct Sign	-0.05	0.874	-0.84*	0.008
	B-E cost	0.47	0.138	-0.30	0.349

* Statistically significant coefficient

Finally, Table 8.27 investigates consistency between forecast measures and economic results. Recall that this Table is directly compared to Table 8.19 for the ASE case. In the LSE case, the PW model exhibits a fully consistent behaviour between forecasts and economic results. The same consistency is also exhibited, but to a lesser degree, by the MA model and by the AR model only with respect to the sign forecast. Hence, in the LSE case (for most models) better forecasts are indeed related to better economic results. However, this relationship is of no practical importance since economic results in the LSE case are non-existing after accounting for transaction costs.

Table 8.27 **LSE:** Spearman rank correlation analysis between forecast accuracy measures and indicators related to economic results.

Model	Indicator	NRMSE	Sign. level	CC	Sign. level
PW	**Correct Sign**	-0.77*	0.014	0.76*	0.016
	B-E cost	-0.84*	0.008	0.84*	0.008
SX	**Correct Sign**	-0.38	0.227	-0.06	0.841
	B-E cost	-0.15	0.625	0.13	0.687
RW	**Correct Sign**	-0.42	0.186	0.40	0.206
	B-E cost	-0.13	0.676	0.11	0.719
MA	**Correct Sign**	-0.58	0.066	0.67*	0.033
	B-E cost	-0.51	0.108	0.71*	0.026
AR	**Correct Sign**	-0.52	0.101	0.61*	0.051
	B-E cost	-0.08	0.796	0.24	0.454

- Statistically Significant coefficient

8.4 CONCLUSIONS

In this Chapter short-term predictions in the ASE and the LSE markets are assessed in terms of forecast accuracy and economic value of the forecasts. The one-day ahead predictions are generated by 5 different linear and nonlinear models. The results are tested for different prediction periods in terms of time length, volatility and sign change frequency characteristics.

The basic qualitative results and conclusions are summarised in Table 9.28. The comparison of the results for the two series reveals significant differences and to a much lesser degree, some similarities between the two markets.

The ASE market is, in general, more predictable than the LSE series by both linear and nonlinear models. This is mainly reflected in the sign prediction indicator and in the economic value of forecasts and, to a lesser degree, in the forecast accuracy measures. Short-term linear dependence seems to exist in both markets. However, it is not sufficiently strong to be economically exploitable after transaction costs.

Nonlinearities found in both markets by the BDS test are also non-exploitable, although in the case of the ASE series, the behaviour of the nonlinear models indicates "indirectly" the possible existence of a chaotic component.

An important differentiating characteristic between the two markets is the existence of "predictability pockets" in the case of the ASE market. Sub-period analysis for the ASE series shows that these periods of increased predictability and economic results (after transaction cost) are characterised by low sign change frequency. On the contrary, volatility does not seem to play an important role in identifying these pockets.

On the other hand, the LSE sub-period analysis has verified the low predictability of this market. The indications of "predictability pockets" were very weak and the only case of economically significant results is due to extraordinary market conditions.

However, the two markets show similar behaviour in terms of the criterion related to increased predictability. So, in the LSE market case, too, lower sign change frequency periods (and not lower volatility periods) are found to be related with periods exhibiting increased predictability.

Table 8.28 Summary of forecast results from the ASE and the LSE series

	ASE SERIES	LSE SERIES
TOTAL PERIOD	➢ Moderate predictability ➢ Linear models marginally superior in terms of forecast measures and economic results. ➢ Nonlinearities non-exploitable in terms of forecast improvement. ➢ Linear or/and nonlinear dependence non-exploitable in economic terms with the exception of the MA model which produces economic results for all different prediction periods. ➢ Indirect indications of chaotic component ➢ Stability of results over different prediction periods.	➢ Low predictability ➢ Nonlinear models marginally superior in forecast measures ➢ Linear or/and nonlinear dependence non-exploitable in economic terms ➢ No indication of chaotic component ➢ Stability of results over different prediction periods.
HIGHEST VOLATIL. PERIOD	➢ High predictability ➢ Nonlinear models superior in all aspects ➢ Most linear and nonlinear forecasts highly exploitable in economic terms	➢ Moderate predictability ➢ Slight superiority of the nonlinear models ➢ Only the MA model produces economic results
LOWEST VOLATIL. PERIOD	➢ Low predictability ➢ Low to marginal superiority of the nonlinear models ➢ No economic value of forecasts	➢ Very low predictability ➢ Marginal superiority of the nonlinear models ➢ No economic value of forecasts
HIGHEST S.C.F. PERIOD	➢ Very low predictability ➢ Low to marginal superiority of the nonlinear models ➢ No economic value of forecasts	➢ Very low predictability ➢ Mixed results regarding models' superiority ➢ No economic value of forecasts
LOWEST S.C.F. PERIOD	➢ High predictability ➢ Nonlinear models superior in all aspects ➢ Most linear and nonlinear forecasts highly exploitable in economic terms	➢ Low predictability ➢ Mixed results regarding models' superiority ➢ No economic value of forecasts

Notice that the level of predictability in the ASE case is not sufficient to challenge market efficiency. This is due not only to theoretical arguments stated by Fama (1991) but also to the fact that "predictability pockets" are not "ex-ante" identifiable[13]. In the opposite case, the market learning process would most probably assimilate the prediction rule and arbitrage should change the law itself. This is most likely in the case that prediction models are very simple, such as naive random walks, moving averages or simple autoregressive specifications.

However, an interesting remark should be made about the ASE market. Recently, the effective transaction cost has dropped to 0.30%-0.35% due to commissions deregulation. With this new threshold, our conclusions for the test periods should be totally different, since economically exploitable results should occur for most of our prediction portfolios (see Figure 8.1g). Lower transaction cost could render the ASE market more exploitable in economic terms, given that its dynamics are similar to the dynamics of the period tested. However in this case too, the argument of an existing arbitrage mechanism, which would eliminate any profit opportunities, remains, so this issue needs further investigation by future research, which should incorporate recent (after 1993) data and could also incorporate more prediction models.

13 Our analysis has revealed a relationship between better predictability and lower s.c.f. However, this relationship was very weak and not statistically significant as far as the Break-even cost indicator is concerned. However, even if this relationship was found to be strong, we would still have to find a way to predict periods with low s.c.f.

Chapter 9

SUMMARY OF RESULTS AND FINAL CONCLUSIONS

The limitations and inadequacies of the traditional linear stochastic framework with respect to explaining the dynamics of the behaviour of financial markets has led the recent research to new nonlinear approaches. Among them, chaotic models have attracted increasing attention since they have been shown to exhibit interesting theoretical and empirical features that could help to gain a better insight of the underlying financial markets mechanisms. However, in order to answer the question of whether low-dimensional chaos may offer a useful way to model financial phenomena, the more fundamental issue of determining whether chaotic behaviour can indeed be observed in financial time series has to be addressed first.

This type of diagnostic analysis has to overcome the lack of a single statistical test for detecting chaos and to face the contradictory results of the existing literature. In this research we have suggested and adopted a "multiple-testing" methodology and we advocate that the combined results of alternative methods and techniques can substantially increase our ability to distinguish between stochastic and chaotic specifications.

We have shown that the bulk of the existing literature from Economics and Finance uses a very limited number of tools, mostly one or two. Moreover, in most cases the methods employed are improperly applied, since standards and precautions set by the literature in the field of the Natural Sciences, where most of these tests have originated from, are not followed.

Hence, the sources of controversies in the literature cannot be fully clarified. When the same data sets are examined in the literature, these controversies might well be due to the inadequacies of the testing framework adopted, including the limited number of tools used and/or problems in applying these methods. However, when different series are examined, we cannot preclude the possibility of different structure, chaotic or not, in different economic and financial series.

In general, this work provides a quite exhaustive analysis of the existing literature, which reveals the inadequacies of the testing framework that has been adopted, so far, either in terms of application problems of specific methods or in terms of the methodology that has been followed. As already mentioned, we advocate for a new methodological approach that definitely relates to the existing framework, since we also employ all the basic methods that can be found in the literature. However, our work adds to it by introducing new methods that have not been used in the Economics and Finance literature, by using additional techniques to increase the power and reliability of existing methods and finally by suggesting a new integrated approach that incorporates all the above.

In empirical terms, this research focuses on the comparison between two financial markets with different levels of maturity: the Athens Stock Exchange, considered to be an emerging market, and London Stock Exchange a mature market in every aspect. This kind of comparison has not been pursued before in the framework of nonlinear dynamics analysis. Our aim was to see whether structural differences exist in terms of chaotic dynamics, which intuitively should be more likely to exist in the case of emerging markets where the degrees of freedom of the underlying system might be fewer. Daily return data, properly validated, was used to describe the dynamics of the two markets, spanning a 13-year and 25-year period for the ASE and the LSE market respectively.

The multiple-testing methodology adopted starts with a statistical analysis in the time and frequency domains, where both series are found to exhibit serial autocorrelation and deviations from normality with fat tails in their distributions.

In the next step, the BDS test, a powerful test for independence, which after proper treatment of the data can also be used as a test for nonlinearity, shows that both series have nonlinear components that are more pronounced for the ASE series, but in both cases they are not due to nonstationarity. This result was compatible with a chaotic explanation so we proceeded further with the R/S test. The latter can reveal long-term memory and fractal structure of the series analysed, even in the presence of noise. In addition, it can distinguish between fractal noise processes and noisy chaos specifications.

For the first time in the relevant literature, our application of the R/S test was enhanced with the bootstrap method for safer and statistically based conclusions and gave the first indications of different structural characteristics in our series. Specifically, the ASE series was found to be noisy, fractal and persistent with an 8-month cycle and a noisy chaos explanation was favoured against a fractal noise one. On the contrary, the LSE series was found to be noisier and no fractality or persistence was possible to be detected.

The chaotic testing framework includes a battery of methods and techniques. Firstly, alternative visual inspection techniques were used, such as the time series plot, the phase-space plot and return maps. All these techniques, which can be proven useful in detecting chaotic structure in the case of pure chaotic systems, were not able to offer much in our case. The plots of both series exhibited increased noise and an obscured structure.

The correlation dimension estimation and related methods and techniques (most of which are firstly applied to financial data), namely Theilers' W specification, the "residual" method, the "wing" and the "shuffle" diagnostics, the "phase randomization" and the "randomized sign" techniques, gave again different results for the two series. The ASE series was found to have a saturating correlation dimension $d \cong 6$, significantly different (lower) than the dimension of the various surrogate series generated by the aforementioned techniques.

The LSE series exhibited non-saturating correlation dimension, which in addition was not significantly different than the dimension of random series having the same distributional characteristics. In both cases statistical significance of our results was assessed through a methodology combining the surrogate data method with the bootstrap method. Moreover, our results were found to be robust to different data lengths through the "independent realizations" method.

The largest Lyapunov exponent estimation was found to verify the possibility of a chaotic component in the ASE series. However, we show that this method, at least with the algorithm that we have employed (the most commonly used in the literature), is not by its own a reliable method to detect chaoticity since it is unable to distinguish between alternative specifications such as chaotic, Gaussian random, and fractal random sequences.

In the next step, SVD analysis was used to reconstruct the phase space and to filter the noise in our two series. We show that this noise-filtering procedure (firstly applied to financial data) is very useful since any existing structure may be fully masked by noise and for this reason it may not be recognizable and detectable by the tools employed.

Phase space reconstruction through SVD analysis verified our previous findings with respect to the dimensionality and the possibility of existence of an attractor for the ASE series but not for the LSE one.
In a second step, the noise-filtered series were analysed further for the existence of a chaotic component through the R/S test, the Correlation dimension estimation and related techniques and the largest Lyapunov exponent estimation. Once more our previous results were verified. The noise-filtered ASE series show the same fractal structure and persistence, a lower correlation dimension of $d \cong 4,5$-5 and the same behavior with respect to the largest Lyapunov exponent estimation.
On the other hand, noise-filtered LSE series show no chaotic structure at all under the same tests. R/S did not show any kind of persistence or fractality and the correlation dimension behaviour was no different than that of the unfiltered series.

In a final step, the existence of a chaotic component in the ASE series was further investigated through nonlinear forecasting techniques. Two nearest neighbour models were employed, the piecewise approximation method and the simplex method. The techniques based on these methods are: the DVS plot, the "varying prediction time" and the "dimensionality" techniques.
The results from these applications were compatible with our previous findings, although in the case of the DVS plot clear results were difficult to be drawn due to inherent shortcomings of this particular test.

In conclusion, none of the different methods and techniques employed gave contradictory results. The two markets were found to be structurally different since for the ASE market the existence of a chaotic component could not be ruled out by all applications. However, this component was found to be mixed with noise, which amounts for more than 60% of the total variance of the series. Hence, the importance of this component is questionable in terms of improved short-term forecasting ability.

In the case of the LSE data, no method could indicate the existence of chaotic structure. On the contrary, all of them were unable to show any evidence against stochasticity and randomness of this data. However, linear dependence in the form of an AR(1) autocorrelation was found, part of which may be due to thin-trading/non-synchronicity bias in the construction of the index as well as some kind of non-linear dependence through the BDS test, which is not of a chaotic nature.

In the last step of our analysis we tried to explore whether the dependencies found in both markets are exploitable in forecasting terms and, furthermore, if these forecasts are of economic importance. This is probably the most interesting part since the real effect of our prior findings is assessed in terms of economic results, the utmost goal as far as financial series are concerned. We simulated an active switching strategy based on return sign forecasts of five different linear and nonlinear models and we compared the economic results with these from a passive buy and hold strategy.

In terms of forecast accuracy, nonlinear models were found unable to exploit the chaotic component in the ASE series. However, forecasting accuracy by all models was better in the case of the ASE series than the LSE one.

In terms of economic results, once transaction costs were taken into account, the LSE market could not be beaten by our forecasts, irrespectively of the forecasting model and prediction period employed, with the exception of the MA model for 1987, which is however the year of the market crash.

The same general conclusion holds for the ASE market, but with some important differences. This market was found to have "predictability pockets", i.e. periods within which positive economic results, after accounting for transaction costs, could be achieved irrespectively of the forecasting model used. Our findings show that the most profitable of these "predictability pockets" are characterized by low sign change frequency and high volatility. However, this finding could not be statistically supported. On the other hand, we found a statistically significant relationship between low sign change frequency periods and increased predictability. Yet, a very weak relationship between the former and periods producing economic results was detected.

In this framework the interesting and rather intriguing finding, regarding the ability of the MA model to produce economic results in the ASE market during all the long prediction periods used, is of limited practical importance. Our analysis shows that the overall performance of the MA model was due to the effect of specific periods, providing "predictability pockets" for this particular model. The occurrence of such pockets, however, is not predictable. Hence, the performance of the MA model does not alter our general conclusion, according to which, the ASE market, although more predictable than the LSE one, does not offer the opportunity of making profits through forecasts, once transaction cost is accounted for. Actually, "predictability pockets", even when they exist, are not "a priori" identifiable. These results, especially with respect to the ASE market, can be considered to support the view that chaos, even if it is present, may be compatible with market efficiency. This is because short-term forecasting improvements chaos may provide are not economically exploitable. However, we should stress again here that our effort in this study is not to challenge market efficiency, not only because chaos can be partly compatible with it, even in theoretical terms, but also because of the problems of such an endeavour [Fama (1991)].

Recapitulating, the results of this research, although they cannot be generalized beyond the specific data sets examined, cast considerable doubt on the importance of chaos theory in explaining the seemingly erratic behavior of the financial markets.
Our findings do not support the generalized use of chaos as an alternative theory behind financial phenomena. We argue that mature markets are highly unlikely to exhibit chaotic components and the empirical findings supporting this view are most probably due to the inadequate testing framework used and to improper application of the methods employed.
The above arguments stem from our findings regarding the LSE market rendering it unpredictable and much more random and noisier than the ASE market. The nonlinear component found by the BDS test is not of a chaotic nature, while the existence of linear dependence can be explained by transaction costs.

On the other hand, our analysis shows that emerging or less developed regional markets, such as the ASE, are likely to exhibit different characteristics than the mature ones in terms of dynamics. Our findings of persistence and long term memory by R/S analysis for the ASE series are consistent with a nonlinear form of lagged dependencies which could be explained by reaction to information in a discontinuous lagged fashion. Based on our experience from the ASE market, we could argue (but not statistically prove) that in less efficient markets, where fundamentals are less reliable and market and pricing models do not offer much help in taking positions, there are investors awaiting for trends to be established before they act.

Our findings with respect to the chaotic component in the ASE market could be related to the fewer number of moving (or controlling) forces prevailing in emerging markets. This is consistent with the increased relative importance of big agents in these markets, who are usually domestic or foreign institutional investors, the latter often acting as market makers in small markets such as the ASE. The role of the State and of big companies with very high market cap should also be stressed in this respect. These active forces could be part of a dynamical system identifiable by the tools employed in this study.

Even if such a system exists, it is highly unlikely to be modelled based on the degrees of freedom corresponding to the correlation dimension found. According to the chaos literature, correlation dimensions greater than 3 make a system, even if chaotic, practically uncontrollable and arbitrarily stochastic (or high-dimensional chaotic since in practice, with the existing tools, it is impossible to distinguish between these two alternatives). Hence, the dimension estimate of the ASE series, which has been found to be approximately 5, is unlikely to be useful for the construction of a chaotic model describing the dynamics of this market.

However, if this information is viewed as a nonlinear analog of an APT framework and of factor analysis, it might be interpreted as giving indications on the number of "risk-factors" priced by the two markets. In this context, it is interesting to contemplate the possibility that, in less sophisticated markets, these factors are fewer (about half according to our results) than the ones in big and mature markets.

In view of the above-stated remarks and from our knowledge of the ASE market, we believe that factors or variables related to marketability, size and some macroeconomic indices could be helpful in a multivariate nonlinear modelling, e.g. through neural nets, and this could be an interesting direction for future research.

In all, chaos, if present in financial markets, is not a widespread phenomenon. Further research should be pursued to decide whether it is more likely to exist and to what extend in emerging and less developed markets. The methodology suggested in this study provides an adequate testing framework for this kind of analysis.
In addition, chaos in financial markets will always be mixed with a great deal of noise, making its identification and exploitation even more difficult. This is clearly reflected to the low ability of nonlinear models to efficiently exploit the chaotic component found in our ASE series.
If we add to the above the unsurpassed data length problem, we should probably put the likelihood of further advances by using this kind of techniques into serious question. Alternatively, future research towards more convincing models of financial markets could be proven much more interesting and fruitful.

However, if someone is interested in putting more effort in the direction of this work, further research could be undertaken towards the direction of exploiting any chaotic or nonlinear components found through alternative nonlinear forecasting models. Our trading simulation model, which is adjustable to include filter rules able to render productive more accurate forecast information (instead of sign changes), provides a very useful tool in accessing the ability of these models to generate economic results.

REFERENCES

Abhyankar, L., Copeland, S. and Wong, W. (1994) "Uncovering nonlinear structure in real-time stock market indexes: The S&P 500, the DAX and the FTSE-100," Working paper # 07/94, University of Stirling.

Abhyankar, L., Copeland, S. and Wong, W. (1997) "Uncovering nonlinear structure in real-time stock market indexes: the S&P 500, the DAX, the Nikkei 225, and the FTSE-100," *Journal of Business and Economic Statistics*, **15**, 1-14.

Abarbanel, H.D.I., Brown, R. and Kadtke, J.B. (1990) " Prediction in chaotic nonlinear systems: Methods for time series with broadbound Fourier spectra," *Physics Review A*, **41**, pp. 1782-1807.

Abraham, N.B., Albano, A.M., Das, B., DeGuzman, G., Yong, S., Gioggia, R.S., Puccioni, G.P. and Tredicce, J.R. (1986) "Calculating the Dimension of Attractor from Small Data Sets" Physics Letters, Vol.114A, **5**, 217-221.

Alexander, S. (1961) "Price Movements in Speculative Markets: Trends or Random Walks," *Industrial Management Review*, **2**, 7-26.

Alogoskoufis, G. and Stengos, T. (1991) "Testing for nonlinear dynamics in historical unemployment series," Discussion Paper, # 1991-4, University of Guelph.

Andrews, D. (1991) "Heteroscedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica,* **59**, pp. 817-858

Anis, A. A. and Lloyd, E.H. (1976) "The Expected Value of the Adjusted Rescaled Hurst Range of Independent Normal Summands," *Biometrica* **63**.

Ariel, R.A. (1987) "A monthly effect in stock returns," *Journal of Financial Economics*, **18**,161-174.

Ariel, R.A. (1990) "High stock returns before holidays: Existence and evidence on possible causes," *Journal of Finance*, **45**, 1611-1626.

Arrow, J.K. (1987) "Workshop on the economy as an evolving complex system: Summary," in *Economy as an Evolving Complex System*, ed. P.W.Anderson, K.J.Arrow and D. Pines.

Ashley, R.D. and Patterson, D.M. (1989) "Linear versus nonlinear macroeconomics: a ststistical test," *International Economic Review*, Vol.30, **3**.

Ashley, R.D., Patterson, D.M. and Hinich, M. (1986) "A Diagnostic Test for Nonlinear Serial Dependence in Time Series Fitting Errors," *Journal of Time Series Analysis*, 7, 165-178.

Atten, P., Caputo, J. G., Malraison, B., Gagne, Y., (1984) "Determination de dimension d' attracteurs pour differents ecoulements", J. Mec. Theor. Application, **133**, (Suppl.)

Aydogan, K. and Booth, G. (1988) "Are there long cycles in common stock returns?," *Southern Economic Journal*, **55**, 141-149.

Badii, R. and Politi, A. (1989) "Dimensions and entropies in chaotic systems," Berlin: Springer Verlag

Badii, R., Broggi, G., Derighetti, B., Ravani, M., Gilberto, S., Politi, A. and Rubio, M.A. (1988) "Dimension increase in filtered chaotic signal," *Physics Review Letters*, **60**, pp. 979-982.

Baek G.E. and Brock W.A (1988) "The Theory of Statistical Inference for nonlinear science: Gauge Functions, Complexity measures and instability measures" Working paper, University of Wisconsin - Madison

Bai-Lin, H. (1984) *Chaos*, Singapore : World Scientific.

Barkoulas, J. and Travlos, N. (1998) "Chaos in an emerging market? The case of the Athens stock exchange," *Applied Financial Economics*, **8**, 231-243.

Barnett, W.A. and Chen P. (1989) "The aggregation theoretic monetary aggregates are chaotic and have strange attractors: an econometric application of mathematical chaos" In W. Barnett, E. Berndt and H. White, (eds.), *Dynamic Econometric Modeling*, Cambridge : Cambridge University Press, pp. 199-246.

Barnett, W.A. and Hinich M.J. (1991) "Has chaos been discovered with economic data?" Working Paper # 153, University of Texas at Austin

Barnett, W.A. , Gallant, A.R., Hinich, M.J. and Jensen, M.J. (1992) "Robustness of nonlinearity and chaos tests to measurement error, inference method and sample size", Discussion paper, Workshop on Nonlinear Dynamics in Economics, European University Institute, Florence , Italy.

Bask, M. (1996) "Dimensions and Lyapunov exponents from exchange rate series," *Chaos, Solitons and Fractals*, 7, 2199-2214.

Bask, M. (1998) "Deterministic chaos in exchange rates?," Working paper, Umea Economic Studies, Umea University, Sweden.

Batterman, R.W. (1993) "Defining Chaos," *Philosophy of Science*, **60**, 43-66.

Baumol, W.J. and Benhabib, J. (1989) "Chaos : Significance, mechanism and economic applications," *Journal of economic Perspectives*, **3**, 77-105.

Beckers, S. (1992) "Trade execution analytics in European equity markets," BARRA Inter-national Ltd.

Benettin, G., Galgani, L., Giorgilli, A. and Strelcyn, J.M. (1980) "Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems: A method for computing all of them," *Meccanica*, **15**, 9-15.

Benhabib, J. and Day, R.H. (1980) "Erratic accumulation," *Economic Letters*, **6**, 113-117.

Benhabib, J. and Day, R.H. (1981) "Rational choice and erratic behavior," *Review of Economic Studies*, Vol. 48, **5**.

Benhabib, J. and Day, R.H. (1982) "A characterization of erratic dynamics in the overlapping generation model," *Journal of Economic Dynamics and Control*, **4**, 27-55.

Blank, S.C. (1991) "Chaos in Futures Markets? nonlinear dynamical analysis," The Journal of Futures Markets, Vol. 11, **6**.

Bodie, Z. (1976) "Common stocks as a hedge against inflation," *Journal of Finance*, **31**, 459-470.

Boldrin, M. (1988) "Persistent oscillations and chaos in dynamic economic models: Notes for a survey," in *The Economy as an Evolving Complex System*, ed. P.W. Anderson, K.J. Arrow and D. Pines, Vol. 5, New York : Addison-Wesley.

Boldrin, M. and Woodford, M. (1990) "Equilibrium models displaying endogenous fluctuations and chaos," *Journal of Monetary Economics*, Vol. 25, **2**.

Bollerslev T. (1986) "Generalized Autoregressive Conditional Heteroskedasticity,", *Journal of Econometrics*, **31**, 307-327.

Bollerslev, T. (1987a) "A conditionally heteroskedastic time series model for speculative prices and rates of return," *The Review of Economics and Statistics*, **69**, 542 - 547.

Bollerslev, T. (1987b) "On the Correlation Structure for the generalized Autoregressive Conditional Heteroskedastic Process," *Journal of Time Series Analysis* Vol.9, **2**, 121 - 131.

Bollerslev, T., Chou, R. and Kroner, K. (1992) "ARCH modelling in finance: A review of the theory and empirical evidence,"

Booth, G. G., Kaen F.R. and Koveos P. (1980) "Currency interdependence in foreign exchange markets," *The Financial Review*, **15**, 36-44.

Booth, G. G., Kaen F.R. and Koveos P. (1982) "R/S analysis of foreign exchange rates under two international monetary regimes," *Journal of Monetary Economics*, **10**, 407-415.

Booth, G. G. and Koveos, P. (1983) "Employment fluctuations : An R/S analysis", *Journal of Regional Science*, **23**, pp. 20-31.

Bountis, T., Karakatsanis, L., Papaioannou, G., and Pavlos, G. (1993) "Determinism and noise in surface temperature time series", *Annales Geophysicae*, **11**, pp. 947-959.

Box, G., and Jenkins, A. (1976) *Time Series Analysis: Forecasting and Control*, Holden Day, Oakland.

Briggs, J. And Peat, F. (1989) *Turbulent Mirror*, New York : Harper and Row.

Brock, W.A. (1986) "Distinguishing random and deterministic systems: Abridged version" *Journal of Economic Theory*, **40**, pp. 168-95.

Brock, W.A. (1987) "Nonlinearity and complex dynamics in Economics and Finance," In the *Economy as an Evolving Complex System*, edited by P.W. Anderson, K.J. Arrow, and D. Pines, Addison-Wesley Publishing Company.

Brock, W.A. and Dechert, W.D. (1987) "Theorems on Distinguishing Deterministic from Random Systems" SSRI Working Paper # 8701, Department of Economics, University of Wisconsin-Madison.

Brock, W.A. and Dechert, W.D. (1990) "Nonlinear dynamical systems: Instability and chaos in economics," *The Handbook of Mathematical Economics*, Vol.4, Ed. W. Hildenbrand and H. Sonnenschein.

Brock, W.A. and DeLima, P. (1995) "Nonlinear time series, complexity theory, and finance," *The Handbook of Statistics*, G.S. Maddala & C.R. Radeds, Vol 14, North Holland. Amsterdam.

Brock, W.A. and Hommes C.H. (1997) "A rational route to randomness," *Econometrica,* **65**, 1059-1095.

Brock, W.A. and Hommes C.H. (1998) "Heterogeneous beliefs and routes to chaos in a simple asset pricing model," *Journal of Economic Dynamics and Control*, 22, 1235-1274.

Brock, W.A. and Malliaris, A. (1989) *Differential Equations, Stability and Chaos in Dynamic Economics,* North Holland.

Brock, W.A. and Sayers, C.L. (1988) "Is the business cycle characterized by deterministic chaos?" *Journal of Monetary Economics,* **22**, pp. 71-90

Brock, W.A., Hsieh, D. and LeBaron, B. (1991) *Nonlinear Dynamics, Chaos and Instability : statistical Theory and Economic Evidence,* MIT Press.

Brock, W.A., Dechert, W.D. and Scheinkman, J. (1987) "A test for independence based on the correlation dimension". SSRI Working Paper # 8702, University of Wisconsin, Madison.

Brock, W.A., Dechert, W.D., Scheinkman, J. and LeBaron, B. (1988) "A test for independence based on the correlation dimension. Manuscript, Madison : University of Wisconsin, Madison.

Brooks, C. (1998) "Chaos in foreign exchange markets: a sceptical view," *Computational Economics*, **11**, pp. 265-281.

Broomhead, D.S. and King, G.P. (1986a) "Extracting qualitative dynamics from experimental data", *Physica D*, **20**, pp. 217-236.

Broomhead, D.S. and King, G.P. (1986b) "On the qualitative analysis of experimental dynamical systems" In S. Sarkar (eds.), *Nonlinear Phenomena and Chaos*, pp. 113-145, Bristol : Adam Hilger.

Broomhead, D.S., Huke, J.P and Muldoon, J.R. (1992) " Linear filters and non-linear systems", *Journal of the Royal Statistics Society B*, **54**, pp. 373-82

Broomhead, D.S., Jones, R. and King, G.P. (1987) "Topological dimension and local co-ordinates from time series data" *Journal of Physics A*, **20**, L563

Campbell, J.Y. and Mankiw, G. (1987) "Are output fluctuations transitory?," Quarterly Journal of Economics, **102**, 857-880.

Campbell, J.Y. and Shiller, R. (1988) "The dividend-price ratio and expectations of future dividends and discount factors," *Review of Financial Studies*, **1**, 195-228.

Caputo, J.G., Malraison, B., Atten, P. (1986) "Determination of attractor dimension and entropy for various flows: An experimentalist's viewpoint," in *Dimensions and Entropies in Chaotic Systems – Quantification of Complex Behaviour*, Springer Verlag in Synergetics Vol.32, Spinger Verlag, Berlin

Carlin, M. (1991) *Neural nets for empirical modelling,* Master's Thesis, Norwegian Institute of Technology (NTH)

Casdagli, M. (1989) "Nonlinear prediction of chaotic time series," *Physica* D, **35**, 335-356.

Casdagli, M. (1991) "Chaos and deterministic versus stochastic nonlinear modeling," *Journal of the Royal Statistics Society*, **54**, No2, 303-328.

Casdagli, M. and Weigend, A. (1993) "Exploring the continuum between deterministic and stochastic modeling," In *Time Series Prediction: Forecasting the Future and Understanding the Past,* Eds. A.S. Weigend and N.A. Gershenfeld, SFI Studies in the Sciences of Complexity, Proc. Vol. XV, Addison-Wesley.

Cecen, A.A. and Erkal, C. (1996) "Distinguishing between stochastic and deterministic behavior in high frequency foreign exchange rate returns: Can nonlinear dynamics help forecasting?," *International Journal of Forecasting,* **12**, 465-473.

Chen, P. (1988) "Empirical and theoretical evidence of economic chaos", *System Dynamics Review,* **4**, pp. 81-108.

Cheung, Y. (1993) "Long memory in foreign exchange rates," *Journal of Business and Economic Statistics,* **11**, 93-101.

Cheung, Y. , Lai, K. and Lai, M. (1993) "Are there long cycles in foreign stock returns?," *Journal of International Financial Markets, Institutions and Money* **3**, 33-47.

Cheung, Y. and Lai, K. (1993) "Do gold markets have long-memory?," *Journal of Business and Economic Statistics* **11**, 93-101.

Chyi, Y-L. (1997) "Nonlinear dynamics and daily stock returns on the Taiwan stock exchange," Applied Financial Economics, 7, 619-634.

Cleveland, W.S. (1979) "Robust locally weighted regression and smoothing scatterplots," *Journal of American Statistical Assosiat ion,* **74 (368)**, 829-836.

Connolly, R.A. (1989) "An examination of the robustness of the weekend effect," *Journal of Financial and Quantitative Analysis*, **24**, 133-69.

Conrand, J. and Kaul, G. (1988) "Time-variation in expected returns," *Journal of Business*, **61**, 409-425.

Cootner, P. (1964) *The Random Nature of Stock Prices*, Cambridge, Mass.: MIT Press.

Copeland, T.E. and Weston, J.F. (1988) *Financial Theory and Corporate Policy*, Addison Wesley.

Cortini, M and Barton, C.C. (1993) "Nonlinear forecasting analyses of inflation-deflation patterns of an active caldera," *Geology*, **21**, 239-242.

Crato, N. (1994) "Some international evidence regarding the stochastic memory of stock returns," Applied Financial Economics, **4**, 281-285.

Cribari-Neto, F. and Zarkos, S.G. (1998) "Bootstrap methods for heteroscedastic regression models: Evidence on estimation and testing," Working paper, University of Illinois at Urbana-Champaign.

Cross, F. (1973) "The behavior of stock prices on Fridays and Mondays," *Financial Analysts Journal*, **29**, 67-79

Cvitanovic, P. (1984) *Universality in Chaos*, Bristol : Adam Hilger

Dana, R.A. and Montruccio, P. (1986) "Dynamic complexity in duopoly games," *Journal of Economic Theory*, Vol. 40, **8**.

Davies, R. and Harte, D. (1987) "Tests for Hurst effect," Biometrica **74**, 95-102.

Day, R.H. (1982) "Irregular growth cycles," *American Economic Review*, **72**, 406-414.

Day, R.H. (1983) "The emergence of chaos from classical economic growth," *Quarterly Journal of Economics*, **98**, 201-213..

Dechert, W.D. (1984) "Does optimal growth preclude chaos? A theorem on monotonicity," *Journal of Economics*, Vol 44,**1**, 57-62.

Dechert, D.W. (1988) "A Characterization of Independence for a Gaussian Process in Terms of the Correlation Integral", Working Paper, # 8812, University of Wisconsin, Madison.

DeCoster, G.P. and Mitchell, D.W. (1991) "Nonlinear monetary dynamics", *Journal of Business and Economic Statistics*, **9**, pp. 455-461.

DeCoster, G.P., Labys, W.C and Mitchell, D.W. (1992) "Evidence of chaos in commodity futures prices", *Journal of Futures Markets*, **12**, pp. 291-305.

DeGrauwe, P. and Vansanten, K. (1990) "Deterministic Chaos in the Foreign Exchange Market," CEPR Discussion paper No. 370.

DeLima, P. (1994) "On the robustness of nonlinearity tests to moment condition failure," Working Papers in Economics No 336, Department of Economics, The Johns Hopkins University.

Deneckere, R. and Pelican, S. (1986) "Competitive Chaos", *Journal of Economic Literature*, Jan, 14-25.

Devaney, R.L. (1989) *An Introduction to Chaotic Dynamical Systems*, Redwood City, CA: Addison-Wesley

Devaney, R.L. (1990) "Chaotic explosions in simple dynamical systems," in *The Ubiquity of Chaos*, ed. Saul Krasner and the American Association for the Advancement of Science.

Diebold, F. and Nason, J. (1990) "Nonparametric exchange rate prediction?," *Journal of International Economics*, **28**, 315-332

Doran, P. (1990) "Search for empirical evidence of strange attractors in historic gold price data", Working paper, # 90/08, University of Essex.

Eckman, J.P. (1981) "Roads to turbulence in dissipative dynamical systems" *Reviews of Modern Physics*, Vol.53, **4**, Part1, 643-654.

Eckman, J.P., Kamphurst, S.O, Ruelle, D. and Gilberto, S. (1988) "Lyapunov exponents from time series" *Physical Review A*, **34**, pp. 4971-4979.

Eckmann, J. P. and Ruelle, D. (1985) "Ergodic theory of chaos and strange attractors," *Review of Modern Physics*, **57**, Vol.3, 617-656.

Eckmann, J. P. and Ruelle, D. (1992) "Fundamental limitations for estimating dimensions and Lyapunov exponents in dynamical systems," *Physica D*, 56, 185-187.

Efron, B. (1979) "Bootstrap methods : Another look at the jack-knife", *The Annals of Statistics*, 7, pp. 1-26.

Efron, B. and Tibshirani, R. (1993) *An Introduction to the Bootstrap*, New York: Chapman and Hall.

El-Gamal, M. (1987) "Simple Methods of Estimation and Inference for Systems Characterized by Deterministic Chaos", University of Rochester, Working Paper No.147.

Elsner. J., and Tsonis, T. (1996) *Singular Spectrum Analysis: A New Tool in Time Series Analysis,* Plenum Press, New York.

Engle, R.F. (1982) "Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation," *Econometrica,* Vol. 50, **4**, 987-1007.

Fama, E. F., and Blume, M. (1966) "Filter Rules and Stock Market Trading Profits," *Journal of Business,* **39**, 226-244.

Fama, E.F. (1965) "The behavior of stock market prices", *Journal of Business,* **38**, pp. 34-105 *

Fama, E.F. (1970a) "Efficient capital markets: Review of theory and empirical work, " *Journal of Finance,* **25** ,383-417.

Fama, E.F. (1970b) "Random Walks in Stock Prices," *Financial Analysts Journal,* **21,** 55-58.

Fama, E.F. (1976) *Foundations of Finance,* New York: Basic Books.

Fama, E.F. (1981) "Stock returns, real activity, inflation and money," *American Economic Review,* **71**, 545-565.

Fama, E.F. (1991) "Efficient capital markets: II," *Journal of Finance,* **46**, 1575-1617.

Fama, E.F. and French, K. (1988) "Permanent and temporary components of stock prices," *Journal of Political Economy,* **96**, 246-273.

Fama, E.F. and Shwert, G.W. (1977) "Asset returns and inflation," Journal of Financial Economics, **5**, 15-146.

Farmer J.D. and Sidorowich J.J. (1987) "Predicting chaotic time series," *Physics Review Letters,* **59**, 845-848.

Feigenbaum, M.J. (1978) "Quantitative universality for a class of nonlinear transformations," *Journal of Statistical Physics,* **19**, 25-52

Feigenbaum, M.J. (1979) "The universal metric properties of nonlinear transformations," *Journal of Statistical Physics,* **21**, 669-706.

Fisher, L. (1966) "Some new stock market Indexes," *Journal of Business,* **39**, 191-225.

Fletcher, R. and Reeves, C.M. (1964) "Function minimization by conjucate gradient," *Computer Journal*, 7, 149-154.

Frank, M.Z. and Stengos, T. (1988) "Some evidence concerning macroeconomic chaos", *Journal of Monetary Economics*, **22**, 423-438.

Frank, M.Z. and Stengos, T. (1989) "Measuring the strangeness of gold and silver rates of return", *Review of Economic Studies*, **56**, 553-567.

Frank, M.Z., Gencay, R. and Stengos, T. (1988c) "International Chaos?", Discussion Paper, # 1988-15, EWP.

Fraser, A.M. (1989a) "Reconstructing attractors from scalar time series: a comparison of singular and redundancy criteria," *Physica D*, **34**

Fraser, A.M. (1989b) "Information and entropy in strange attractors," *IEEE Transactions on Information Theory*, Vol. 35, **2**.

Fraser, A.M. and Swinney, H.L. (1986) "Independent coordinates for strange attractors from mutual information," *Physical Review A*, Vol.33, **2**.

French, K. R. (1980) "Stock Returns and the Weekend Effect," *Journal of Financial Economics*,

Gershenfeld, N. (1987) "An Experimentalist's Introduction to the Observation of Dynamical Systems" *Directions in Chaos*, Vol 2, 310-383.

Gershenfeld, N. (1992) "Dimension measurement of high-dimensional systems," *Physica D*, **8**,135-154

Ghil, M. and Mo, K. (1991) "Interseasonal Oscillations in the global atmosphere. Part I: Northern hemisphere and tropics," *Journal of Atmospheric Sciences*, **5**, 752-779.

Gibbons, M.R. and Hess, P. (1981) "Day of the Week Effects and Asset Returns," *Journal of Business*, **54**, 3-27.

Gibson, J.F., Farmer, J.D., Casdagli, M. and Eubank, S. (1992) "An analytic approach to practical state space reconstruction," *Physica D*, **57**, 1-30.

Gleick, J. (1990) *Chaos: Making a New Science*, New York : Viking Press.

Godfrey, M..D., Granger, J. and Morgenstern, O. (1964) "The Random Walk Hypothesis of Stock Market Behavior," *Kyklos*, **17**, 1-30.

Gollub, J.P., Romer, E.J. and Socolar, J.E. (1980) "Trajectory divergence for coupled relaxation ocillators: measurement for geostrophic turbulence," *Journal of Statistical Physics*, 23, 321-334.

Grandmont, J. (1985) "On endogenous competitive business cycles," *Econometrica*, **53**,1985.

Granger, C.W.J. (1972) "Empirical studies of capital markets: A survey," in *Mathematical Methods in Investment and Finance*, editors G. Szego and K. Shell, North-Holland, Amsterdam.

Granger, C.W.J. and Joyeux, R. (1980) "An introduction to long-range time series models and fractional differencing," *Journal of Time Series Analysis,* **1**,15-30.

Granger, C.W.J., and Anderson, A.P. (1978) *An Introduction to Bilinear Time Series Models*, Gottingen: Vandenboech & Ruprecht.

Grassberger, P. (1986) "Do climatic attractors exist?," Nature, **323**, 609-615.

Grassberger, P. (1988) "Finite Sample Corrections to Entropy and Dimension Estimates," *Physics Letters A*, Vol 128, No 6,7 , pp. 369 - 373.

Grassberger, P. and Procaccia, I. (1983a) "Characterization of strange attractors", *Physics Review Letters*, **50**, pp. *

Grassberger, P. and Procaccia, I. (1983b) "Measuring the strangeness of strange attractors", *Physica D*, **9**, pp. 189-208.*

Grebogi, C., Ott, E. Pelican, S. and Yorke, J.A. (1984) "Stange attractors that are not chaotic, " *Physica D*, **13**, 261-268.

Greene, M. and Fielitz, B. (1977) "Long-term dependence in common stock returns," *Journal of Finencial Economics* **4**, 339-349.

Guckenheimer, J. (1984) "Dimension estimates for attractors" Contemporary Mathematics, 28, 357 -371.

Gutelkin, M.N. and Gutelkin, B.N. (1983) "Stock Market Seasonality: International Evidence," *Journal of Financial Economics*, 34, 122-154..

Scienze Economiche e Commerciali, Vol.36, **3**.

Harris, L. (1986) "A transaction data study of weekly and intradaily patterns in stock returns," *Journal of Financial Economics*, **16**, 99-117.

Haubrich, J.G. and Lo, A.W. (1989) "The sources and nature of long-term memory in business cycle" Working Paper, # 5-89, Rodney L. White Center for Financial research, The Wharton School, University of Pennsylvania, Philadelphia.

Haugen, R. A. and Lakonishok, J. (1988) *The Incredible January Effect*, Homewood : Dow Jones-Irwin.

Henon, M. (1976) "A two-dimensional mapping with a strange attractor," *Communications in Mathematical Physics*, **50**.

Hiemstra, G. and Jones J.D. (1995) "Another Look at Long Memory in Common Stock Returns", Discussion Paper, University of Strathclyde.

Hinich, M. J. (1982) "Testing for Gaussianity and linearity of a stationary time series," Journal of Time Series Analysis, Vol.3, **3**,169-76.

Hinish, M. and Patterson, D. (1985) "Evidence of nonlinearity in stock returns," *Journal of Business and Economic Statistics* **3**, 69-77.

Hinish, M. and Patterson, D. (1988) "Evidence of nonlinearity in the trade-by-trade stock market return generating process," in *Economic Complexity: Chaos, Sunspots, Bubbles and Nonlinearity*, ed. W.A.Barnett, J. Geweke, K. Shell.

Hollander, M. and Wolfe, D.A. (1973) *Nonparametric Statistical Methods*, New York: John Wiley and Sons, Inc.

Hosking, J. (1981) "Fractional differencing," *Biometrika*, **68**, 165-176.

Horowitz, J.L. (1997) "Bootstrap methoda in econometrics: theory and numerical performance," In Kreps, D.M. & Wallis, K.F., eds., *Advances in Economics and Econometrics: Theory and Applications* (Seventh World Congress), vol.3, pp. 188-222. New York : Cambridge University Press.

Hsieh, D.A. (1989) "Testing for nonlinear dependence in daily foreign exchange rates: 1974-1983", *Journal of Business*, **62**, pp. 339-66.

Hsieh, D.A. (1991) "Chaos and nonlinear dynamics: application to financial markets", *Journal of Finance*, **46**, pp. 1839-77.

Hsieh, D.A., and LeBaron, B. (1988) "Finite sample properties of the BDS statistic I: Distribution under the null hypothesis", Manuscript, Chicago: University of Chicago.

Hsu, D., Miller, R. and Wichern, D. (1974) "On the stable paretian behavior of stock-market prices," *Journal of the American Statistical Association*, **69**, 108-113.

Hudson, J.L. and Mankin, J.C. (1981) "Chaos in the Belousov-Zhabotinski reaction," *Journal of Chemical Physics*, 74, 6171-6177.

Hunter, N.F.Jr. (1992) "Nonlinear prediction of speech signals," In M. Casdagli and S. Eubanks, editors, *Nonlinear Modeling and Forecasting*, Addison-Wesley, 467-492.

Hurst, H.E. (1951) "Long-term storage of reservoirs", *Transactions of the American Society of Civil Engineers*, **116**.

Jensen, M.C. and Bennington, G.A. (1970) "Random walks and technical theories: Some additional evidence," Journal of Finance, 5.

Jeong, J. and Maddala, G.S. (1993) "A perspective on application of bootstrap methods in econometrics," In Maddala, G.S., Rao, C.R. & Vinod, H.D., eds., *Handbook of Statistics: Econometrics*, vol. 11, pp. 573-610. Amstardam: North-Holland.

Johnson, R.A., Palmer, K.J. and Sell, G. (1984) "Ergodic properties of linear dynamical systems," Preprint, Minneapolis.

Kahneman, D. and Tversky, A. (1979) "Prospect theory: an analysis of decision under risk," *Econometrica*, XLVII, 263-291.

Kahneman, D. and Tversky, A. (1984) "Choises, Values and Frames,"*American Psycologist*, XXXIX, 341-350.

Kao, G. W. and Ma, C.K. (1992) "Memories, heteroskedasticity and price limit in currency futures markets," *The Journal of Futures Markets*, Vol.12, **6**, 679-692.

Kato, K. and Schallheim, J. (1985) "Seasonal and Size Anomalies in the Japanese Stock Market," *Journal of Financial and Quantitative Analysis*, **11**. 254-263.

Keim, D. and Stambauch, R. (1984) "A Further Investigation of the Weekend Effect in Stock Returns," *Journal of Finance*, **22**, 85-94..

Keim, D.B. (1983) "Size-Related Anomalies and Stock Return Seasonality: Further Empirical Evidence," *Journal of Financial Economics*, **12**, 13-32.

Keim, D.B. (1985) "Dividend Yields and Stock Returns: Implications of Abnormal January Returns," *Journal of Financial Economics*, **39**, 27-39..

Kelsey D. (1988) "The Economics of Chaos or the Chaos of Economics," *Oxford Economic Papers*, **40**, 1-31.

Kennel, M.B. and Isabelle, S. (1992) "Method to distinguish possible chaos from colored noise and to determine embedding parameters," *Physical Review A*, **46** (6), 3111-3118.

Kolmogorov, A.N. (1961) "Local structure of turbulence in an incompressible liquid for very large Reynolds numbers," in *Turbulence : Classic Papers on statistical Theory*, ed. S.K. Friedlander and L. Topper, New York : Interscience.

Kostelich, J.E. (1992) "Problems in estimating dynamics from data," *Physica D*, **58**, 138-152.

Kostelich, J.E. and Yorke, A.J. (1990) "Noise reduction: Finding the simplest dynamical system consistent with the data," Physica D, **41**, 183-196.

Krishna Mohan, T.R. and Subba Rao, J. (1989) "Dimension analysis of climatic data," *Journal of Climate*, Vol.2, **9**, 1047-1057.

Kugiumtzis, D. and Christophersen N. (1994) "State space reconstruction : Method of Delays vs. Singular Spectrum Approach," Preprint, Dept. of Informatics, University of Oslo.

Kuhn, T.S. (1962) *The Structure of Scientific Revolutions*, Chicago : University of Chicago Press.

Kunsch, H. R. (1989) "The Jacknife and the Bootstrap for general stationary observations," *The Annals of Statistics*, **17**,1217-1241.

Lakonishok, J. and Smidt, S. (1988) "Are seasonal anomalies real?: A ninety year perpective," *Review of Financial Studies*, **1**, 435-455.

Lapedes, A. and Farber, R. (1987) "How newral nets work," In D.Z. Anderson, editor, Neural Information Processing Systems, American Institute of Physics, New York, 442-456.

Larrain, M. (1991a) "Empirical tests of chaotic behavior in a nonlinear interest rate model," *Financial Analysts Journal*, **47**.

Larrain, M. (1991b) "Testing chaos and nonlinearities in T-Bill rates," *Financial Analysts Journal*, September/October, 51-62.

Larsen, E.R., Mosekilde, E. and Sterman, J.D. (1990) "Chaotic dynamics produced by human decision making behaviour," Working parer, 1-90, Institute of Economics : Copenhagen Business School.

LeBaron, B. (1991) "The changing structure of stock returns," In *Nonlinear Dynamics, Chaos and Instability*, edited by W.A. Brock, D. Hsieh and B. LeBaron, Cambridge, MA: MIT Press.

LeBaron, B. (1992a) "Nonlinear forecasts for the S&P Stock Index" Nonlinear Modeling and Forecasting, SFI Studies in the Science of Complexity, Proc. Vol. XII, Eds. M. Casdagli & S. Eubank, Addison-Wesley.

LeBaron, B. (1992b) "Forecast improvements using a volatility index," *Journal of Applied Econometrics*, 7, 137-149.

Lee, T-H., White, H. and Granger, C.W.J. (1993) "Testing for neglected nonlinearity in time series models: A comparison of neural network methods and alternative tests," *Journal of Econometrics*, **56**, 269-290.

Li, T-Y. and Yorke, J. A. (1975) "Period three implies chaos," *American Mathematical Monthly*, Dec., 985-992.

Linden, N., Satchell, S. and Yoon, Y. (1992) " Predicting Britich financial Idices : An approach based on chaos theory", Discussion paper, Workshop on Nonlinear Dynamics in Economics, European University Institute, Florence, Italy.

Liu R.Y. and Singh, K. (1992) "Moving blocks Jacknife and Bootstrap capture weak dependence," In *Exploring the Limits of Bootstrap* by LePage R. and Billard L. eds. John Wiley and Sons, 225-248.

Liu, T., Granger, C. and Heller, W. (1992) "Using the correlation exponent to decide if an economic series is chaotic," *Journal of Applied Econometrics*, 7, 25-39.

Lo, A. and Mackinlay, A.C. (1988) "Stock market prices do not follow random walks: evidence from a simple spscification test," *Review of Financial Studies*, **1**, 41-66.

Lo, A. W. (1991) "Long term memory in stock market prices," *Econometrica*, **59**, 1279-1313.

Lorenz, E. N. (1963) "Deterministic nonperiodic flow," *Journal of Atmospheric Sciences*, **20**, 130-141

Lorenz, H. W. (1987a) "International trade and the possible occurrence of chaos," *Economics Letters*, **23**, 135-138.

Lorenz, H. W. (1987b) "Strange attractors in a multisector business cycle model," *Journal of Economic Behavior and Organization*, **8**, 397-411.

Loteran, M. and Phillips, P. (1994) "Testing the covariance stationarity of heavy tailed time series: An overview of the theory with applications to several financial data sets," *Journal of Empirical Finance*, **1**, 211-248.

Mackey, M. C. and Glass, L. (1977) "Oscillations and chaos in physiological control systems," *Science*, **197**, 287-289.

Malkiel, B. (1981) *A Random Walk Down Wall Street*, New York: W.W. Norton.

Malliaris, G.A. and Philippatos, G. (1992) "Random walk vs. chaotic dynamics in financial economics," Proceedings of the European Conference on Chaos in the Social Sciences, Springer-Verlag.

Mandelbrot, B. (1963) "The variation of certain speculative prices," *Journal of Business*, **36**, 394-419.

Mandelbrot, B. (1966) "The variation of some other speculative prices," *Journal of Business*, **39**, 223-238.

Mandelbrot, B. (1972) "Statistical methodology for non-periodic cycles: from the covariance to R/S analysis", *Annals of Economic and Social Measurement*, **1**, 1972.

Mandelbrot, B. (1982) *The Fractal geometry of Nature*, New York : W.H. Freeman .

Mandelbrot, B. and J. Wallis (1969a) "Computer experiments with fractional Gaussian noises. Part 1, averages and variances", *Water Resourses Research*, **5**.

Mandelbrot, B. and Wallis, J. (1969b) "Computer experiments with fractional Gaussian noises Part 2, rescaled ranges and spectra", *Water Resourses Research*, 5.

Mandelbrot, B. and Wallis, J. (1969c) "Computer experiments with fractional Gaussian noises. Part 3, mathematical appendix", *Water Resourses Research*, **5**.

Mandelbrot, B. and Wallis, J. (1969d) "Robustness of the rescaled range R/S in the measurement of non-cyclic long run statistical dependence", *Water Resourses Research*, **5**.

May, R.M. (1976) "Simple mathematical models with very complicated dynamics," *Nature*, **261**, 459-467.

Mayer-Kress, G. (1987) "Application of Dimension Algorithms to Experimental Chaos," in *Directions in Chaos*, Ed. Hao Bai-Lin, World Scientific, Singapore.

Mayfield, E.S. and Mizrach, B. (1992) "On determining the dimension of real-time stock-price data," *Journal of Business & Economic Statistics*, **10**, 367-74.

Medio, A. (1992) "Chaotic dynamics : theory and applications to economics", Cambridge: Cambridge University Press.

Mees, A.I., Rapp, P.E. and Jennings, L.S. (1987) "Singular value decomposition and embedding dimension," *Physical Review A*, Vol. 36.

Meese, R.A. and Rose, A.K. (1989) "An empirical assessment of nonlinearities in models of exchange rate determination," Technical paper, University of California, Berkeley.

Mills, T. (1992) "How robust is the finding that innovations to U.K. output are persistent ?," *Scottish Journal of political Economy*, **39**, 154-166.

Mills, T. (1993) "Is there Long-term Memory in UK Stock Returns?," *Applied Financial Economics*, **3**, 303-306.

Milonas, N., Koveos, P. And Booth, G. (1985) "Memory in commodity futures contracts: A comment, " *The Journal of Futures Markets*, **5**, 113-114.

Mitschke, F., Moller, M. and Lange, W. (1988) "Measuring filtered chaotic signals", *Physics Review A*, **37**, 4518-4521.

Mizrach, B. (1989) "Multivariate nearest-neighbor forecasts of EMS exchange rates," Technical paper, Boston University, Chestnut Hill.

Nagashima, H. (1982) "Experiment on chaotic response of forced Belousov-Zhabotinski reaction," *Journal of Physics Society Japan*, **51**, 21-29.

Nelson, C.R. and Plosser, C.I. (1982) "Trends and random walks in macroeconomic time series," Journal of Monetary Economics, **10**, 139-162.

Nelson, C.R. and Startz, R. (1990a) "The distribution of the instrumental variable estimator and its t ratio when the instrument is a poor one," *Journal of Business*, **63**, 125-140.

Nelson, C.R. and Startz, R. (1990b) "Some further results on the exact small sample properties of the instrumental variable estimator," *Econometrica*, **55**, 703-8.

Nerenberg, M.A. and Essex, C. (1990) " Correlation dimension and systematic geometric effects", *Physics Review A*, **42**, pp. 7065-7074.

Noreen, E.W., (1989) "Computer Intensive Methods for Testing Hypotheses," New York: John Wiley & Sons.

Osborne, A.R. and Provenzale, A. (1989) "Finite Correlation dimension for stochastic systems with power-law spectra," *Physica D*, **35**, pp. 357-381.

Oseledec, V.I. (1968) "A multiplicative ergodic theorem and Lyapunov characteristic numbers of dynamical systems," *Transactions of the Moscow Mathematical Society*, **19**, 197-231.

Ott, E. (1981) "Strange attractors and Chaotic motions of dynamical systems," *Reviews of Modern Physics*, Vol53, **4**, Part1, 655-671.

Ott, E., Grebogi, C., and Yorke J.A. (1990) "Controlling Chaos", *Physical Review Letters*, **64**, 1196.

Packard, N.H., Crutchfield, J.P., Farmer, J.D. and Shaw, R.S. (1980) "Geometry from a time series," *Physical Review Letters*, Vol. 45, **9**.

Pagan, A. and Schwert, G. (1990) "Testing for covariance stationarity in stock market data," *Economics Letters*, **33**, 165-170.

Panas, E. and Nini, V. (1998) "Are oil markets chaotic? A nonlinear dynamics analysis," Working paper, Athens University of Economics and Business.

Pavlos, G.P., Rigas, A.G., Dialetis, D., Sarris, E.T.,Karakatsanis, L.P., Tsonis, A.A. (1992a) "Evidence for chaotic dynamics in the outer solar plasma and the earth magnetosphere," *Chaotic Dynamics: Theory and Practice* Ed. T. Bountis, Plenum Press, New York.

Pavlos, G.P., Kyriakou, G.A., Rigas, A.G., Liatsis, P.I., Trochoutsos, P.C. and Tsonis, A.A. (1992b) "Evidence for strange attractor structures in space plasmas," *Annales Geophysicae*, **10**, 309-322.

Pavlos, G.P., Karakatsanis, L. Latoussakis, J.B., Dialetis, D. and Papaioannou G. (1994) "Chaotic analysis of a time series composed of seismic events recorded in Japan," *International Journal of Bifurcation and Chaos*, Vol.4., **1**, 87-98.

Pawelzik, K. and Schuster, H.G. (1991) "Unstable periodic orbits and prediction," *Physics Review A*, **43**, 1808-1812.

Pesaran, M.H. and Timermann, A.G. (1992) "The statistical and economic significance of the predictability of excess returns on common stocks," DAE Working paper No 9022.

Peters, E. (1989) "Fractal structure in the capital markets", *Financial Analysts Journal*, July/August.

Peters, E. (1991a) "Chaos and order in the capital markets: a new view of cycles, prices and market volatility", New York: John Wiley & Sons.

Peters, E. (1991b) "A chaotic attractor for the S&P 500", *Financial Analysts Journal,* March/April.

Peters, E. (1992) "R/S analysis using logarithmic returns : a technical note", *Financial Analysts Journal,* November/December.

Peters, E. (1994) "Fractal market analysis : applying chaos theory to investment and economics," New York: John Wiley & Sons.

Pilarinu, E. (1993) "A chaotic approach to the analysis of financial data", Technical paper, The University of Tennesee, Knoxville.

Pilgram, B., Schappacher, W. and Pfurtsceller, G. (1992) "A noise reduction method using singular value decomposition," Working Paper, University of Graz, Austria.

Ploeg, F. Van der (1985) "Rational expectations, risk and chaos in financial markets," *The Economic Journal,* 7.

Pohjola, M. (1985) "Built-in flexibility of progressive taxation and the dynamics of income: Stability, cycles or chaos?," *Public Finance,* Vol. 40, **2**, 263-271.

Poterba, J. and Summers, L. (1988) "Mean reversion in stock returns : evidence and impications," *Journal of Financial Economics,***22**, 27-60.

Praschnik, J. (1991) "Non-linear dependence? A look at the treasury bill futures market", *Applied Financial Economics,* **1**, pp. 165-73.

Prescott, D.M. and Stengos, T. (1988) "Do asset markets overlook exploitable nonlinearities? The case of gold," Technical paper, University of Guelph, Ontario.

Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1988) *Numerical Recipes in C,* Cambridge: Cambridge University Press

Priestley, M.B. (1981) *Spectral Analysis and Time Series,* Academic Press, London.

Prigogine, I. (1980) *From Being to Becoming,* New York : W.H. Freeman

Prigogine, I. , and Stengers, I. (1984) *Order Out of Chaos,* New York : Bantan Books

Provenzale, A., Smith L.A., Vio, R., and Murante, G. (1992) "Distinguishing low-dimensional dynamics and randomness in measured time series," *Physica D,* **58**, 31-49.

Ramsey, J.B. and Yuan, H.J. (1989) "Bias and error bars in dimension calculations and their evaluation in some simple models," *Physics Letters A*, **134**, pp. 397-98

Ramsey, J.B. and Yuan, H.J. (1990) "The statistical properties of dimension calculations using small data sets," *Nonlinearity*, **3**, 155-176.

Ramsey, J.B. Sayers, C.L. and Rothman, P. (1990) "The statistical properties of dimension calculations using small data sets: some economic applications," *International Economic Review*, **31**, 991-1020.

Rosenblatt, M. (1956) "A central limit theorem and a strong mixing condition," *Proceedings of the National Academy of Sciences* **42**, 43-47.

Rossler, O.E. (1976) "An equation for continuous chaos," *Physics Letters A*, **57**, 397-403.

Rozeff, M.S. (1984) "Dividend yields are equity risk premiums," *Journal of Portfolio Management*, **11**, 68-75.

Rozeff, M.S., and Kinney, W.R. (1976) "Capital Market Seasonality: The Case of Stock Returns," *Journal of Financial Economics*, **11**.

Ruelle, D. (1983) "Small random perurbations of dynamical systems and the definition of attractors," *Commun. Mathematical Physics*, **82**,137-151.

Ruelle, D. (1984) "Characteristic exponents for a viscous fluid subjected to time dependent forces," *Commun. Mathematical Physics*, **93**, 285-297

Ruelle, D. (1990) "Deterministic Chaos: the science and the fiction" *Proceedings of the Royal Society of London A*, **427**, 241-248

Samuelson, P.A. (1965) "Proof That Properly Anticipated Prices Fluctuate Randomly," *Industrial Management Review*, **6**, 41-49.

Satchell, S. and Timmermann A. (1992a) "An assessment of the economic value of nonlinear foreign exchange rate forecasts," Discussion paper in Financial Economics, FE-6/92, Birbeck College, University of London.

Satchell, S. and Timmermann A. (1992b) "Daily returns in European stock markets: Predictability, Nonlinearity and Transaction Costs," Discussion paper in Financial Economics, FE-5/92, Birbeck College, University of London.

Sauer, T. (1992) "A noise reduction method for signals from nonlinear systems," *Physica D*, **58**, 193-201.

Savit, R. (1988) "When random is not random: An introduction to chaos in market prices," *The Journal of Futures Markets*, Vol. 8, **3**, 271-289.

Savit, R. (1989) "Nonlinearities and Chaotic effects in options prices", *Journal of Futures Markets*, **9**, 507-518.

Schuster, H.G. (1988) *Deterministic Chaos - An Introduction*, Weinheim: Physik-Verlag.

Schwartz, G. (1978) "Estimating the dimension of a model" *Annals of Statistics* **6**, 461-464.

Scott, D.W. (1992) *Multivariate Density estimation,* Wiley, New York.

Sewel, S.P., Stansell, S.R., Lee, I. and Below, S.D. (1996) "Using chaos maesures to examine international capital market integration," *Applied Financial Economics*, **6**, 91-101.

Shaffer, S. (1991) "Structural Shifts and the Volatility of Chaotic Markets," *Journal of Economic Behavior and Organization*, **15**, 201-214.

Shannon, C.E. and Weaver, W. (1963) *The Mathematical Theory of Communication,* Urbana : University of Illinois.

Shaw, R.S. (1981) "Strange attractors, chaotic behavior and information flow," *Z. Naturforsh*, **36a**, 80-90.

Sheinkman, J. and LeBaron, B. (1989) "Nonlinear dynamics and GNP data", In W. Barnett, J. Geweke, and K. Shell (eds.), *Economic Complexity: Chaos, Sunspots, Bubbles, and Nonlinearity*, Cambridge: Cambridge University Press.

Sheinkman, J. and LeBaron, B. (1989) "Nonlinear Dynamics and Stock Returns", *Journal of Business*, **62**, 311-337.

Shiller, R.J. (1981) "Do stock prices move too much to be justified by subsequent changes in dividends," *American Economic Review*, **71**, 421-436.

Shiller, R.J. (1989) *Market Volatility*, Cambridge, MA : MIT Press.

Shimada, I. and Nagashima, T. (1979) "A numerical approach to ergodic problem of dissipative dynamical systems," *Prog. Theor. Phys.*,**61**, 1605-1622.

Smale, S. (1963) "Diffeomorphisms with many periodic points," in *Differential and Combinatorial Topology*, Ed. S.S. Cairns, Princeton : Prinnceton University Press.

Smith, A.L. (1988) "Intrinsic limits on dimension calculations," *Physics Letters A*, **133**, 283-288.

Smith, A.L. (1992) "Estimating dimension in noisy chaotic time series," *Journal of the Royal statistical Society*, **54**.

Smith, S. W. and Hoy, J.L. (1990) "Chaos in Economics The Case of a Stock Market Crash," The Polytechnic of North London, Working Paper No. 34.

Stutzer, M. (1980) "Chaotic dynamics and bifurcationsin a macro model," *Journal of Economic Dynamics and Control*, **2**, 353-376.

Sugihara, G. and May, R.M. (1990) "Nonlinear forecasting as a way of dinstiguishing chaos from measurement error in time series," *Nature*, **344**, 734-41.

Swinney, H.L. (1983) "Observation of order and Chaos in Nonlinear Systems," *Physica D*, 7, 3-15

Takala, K. and Viren, M. (1996) "Chaos and nonlinear dynamics in financial and nonfinancial time series: Evidence from Finland," *European Journal of Operational Research*, **93**, 155-172.

Takens, F. (1980) "Detecting strange attractors in fluid turbulence" in *Dynamical Systems and Turbulence*, Ed. D. Rand and L.S. Young, Lecture Notes in Mathematics, Vol.898, 366-382, Springer-Verlag, Berlin.

Takens, F. (1983a) "Distinguishing deterministic and random systems", In G. Borenblatt, G. Iooss, and D. Joseph (eds.), *Nonlinear Dynamics and Turbulence*, Boston: Pitman Advanced.

Takens, F. (1983b) "Invariants related to dimension and entropy" Proceedings of the Thirteenth Coloquio Brasileiro de Matematica. Rio de Janeiro: Instituto de Matematica Pura e Apicada.

Takens, F. (1984) "On the numerical determination of the dimension of an attractor," in *Dynamical Systems and Bifurcations*, Vol. 1125, Lecture Notes in Mathematics, Springer-Verlag, 99-106.

Tata, F. and Vassilicos, C. (1991) " Is There Chaos in Economic Time Series? A Study of the Stock and the Foreign Exchange Markets", LSE Financial Markets Group, Working Paper # 120, London School of Economics.

Theiler, J. (1986) "Spurious dimension from correlation algorithms applied to limited time-series data", *Physical Review A*, **34**, 2427-32.

Theiler, J. (1990a) "Estimating fractal dimension," *Journal of Optical Society of America*, Vol.7, **6**, 1055-1073.

Theiler, J. (1990b) "Statistical precision of dimension estimators," *Physics Review A*, **41**, 3038-3051.

Theiler, J. (1991) "Some comments on the correlation dimension of $1/f^a$ noise", *Physics Letters A*, **155**, pp. 480-493 1990

Theiler, J. and Eubank, S. (1993) "Don't bleach chaotic data," *Chaos*, **3**,771-782.

Theiler, J., Eubank, S. Longtin, A. Galdrikian, B. and Farmer J.D. (1992) "Testing for nonlinearity in time series: the method of surrogate data", *Physica D*, **58**, 77-93.

Timmermann, A. and Satchell, S. (1992) "An international comparison of non-linear patterns in daily stock returns", Discussion paper, Workshop on Nonlinear Dynamics in Economics, European University Institute, Florence , Italy.

Tong, H. (1992) "Some comments on a bridge between nonlinear dynamicists and statisticians", *Physica D*, **58**, 200-303.

Tong, H., and Lim, K.S. (1980) "Threshold autoregression, limit cycles and cyclical data," *Journal of the Royal Statistical Society*, **42B**, 245-292.

Townshend, B. (1992) "Nonlinear prediction of speech signals," In M. Casdagli and S. Eubanks, editors, *Nonlinear Modeling and Forecasting*, Addison-Wesley, 433-453.

Tsay, R. (1986) "Nonlinerity tests for time series," *Biometrica*, **73**, 23-31.

Tsonis, A.A. (1992) *Chaos: From Theory to Applications*, Plenum Press, New York

Tsonis, A.A. and Elsner, J.B. (1992) "Nonlinear prediction as a way of distinguishing chaos from random fractal sequences", *Nature*, **358**, 217-220.

Tsonis, A.A., J.B. Elsner and K.P.Georgakatos (1992) "Estimating the dimension of weather and climate attractors: important issues about the procedure and interpretation", Paper submitted to the *Journal of Atmospheric Sciences*.

Turner, A.L. and Weigel, E.J. (1990) "An analysis of stock market volatility," *Russell Research Commentaries*, Frank Russell Co. Tacoma, WA.

Tversky, A. (1990) "The psychology of risk," in *Quantifying the Market Risk Premium Phenomena for Investment Decision Making*, Charlottesville, VA: Institute of Chartered Financial Analysts.

Tversky, A. and Kahneman, D. (1974) "Judgement under uncertainty : heuristics and biases," *Science*, **185**, 1124-1131.

Tversky, A. and Kahneman, D. (1991) "Loss aversion in riskless choice: a reference-dependent model," *The Quarterly Journal of Economics*, November, 1039-1061.

Vaga, T. (1991) "The coherent market hypothesis," *Financial Analysts Journal*, December/January.

Vaidyanathan, R. & Krehbiel, T. (1993) "Does the S&P 500 futures misprising series exhibit nonlinear dependence across time ?", *The Journal of Futures Markets*,

Vassilicos, J.C. (1990) "Are Financial Markets Chaotic?: A Preliminary Study of the Foreign Exchange Market", LSE Financial Markets Group, Discussion Paper No 86, London School of Economics.

Vassilicos, J.C., Demos, A. and Tata, F. (1992) "No evidence of chaos but some evidence of multifractals in the foreign exchange and the stock markets," LSE Financial markets Group Discussion Paper Series.

Vautard, R. and Ghil, M. (1989) "Singular spectrum analysis in nonlinear dynamics with applications to paleoclimatic time series", *Physica D*, **35**, 395-424.

Vautard, R., Yiou, P. and Ghil, M. (1992) "Singular spectrum analysis: A toolkit for short, noisy chaotic signals," *Physica D*, **58**, 95-126.

Wallis J. and N. Matalas (1970) "Small sample properties of H and K estimators of the Hurst coefficient h," *Water Resources Research*, **6**, 332.

Weigend, A., Huberman, B. and Rumelhart, D.E. (1990) "Predicting the future: a connectionist approach," *International Journal of Neural Systems*, **1**, 193-209.

Weigend, A., Huberman, B. and Rumelhart, D.E. (1992) "Predicting sunspots and exchange rates with connectionist networks," In M. Casdagli and S. Eubank, editors, *Nonlinear Modeling and Forecasting* ,Addison-Wesley, 395-432.

White, H. (1988) "Economic prediction using neural networks: The case of IBM daily stock returns," Technical paper, University of California, San Diego.

Wiggins, S. (1990) *Introduction to Applied Nonlinear Dynamical Systems and Chaos,* **New York, Springer Verlag.**

Willey, T. (1992) "Testing for nonlinear dependence in daily stock indices", *Journal of Economics and Business*, **44**.

Winnie, J.A. (1992) "Computable Chaos," *Philosophy of Science*, **59**, 263-275.

Wolf, A. and Swift, J. (1984) "Progress in computing Lyapunov Exponents from experimental data," in Statistical Physics and Chaos in Fusion Plasmas, Ed. C.W. Horton Jr.and L.E. Reichl, Wiley, New York.

Wolf, A., J.B. Swift, H.L. Swinney and J. Vastano (1985) "Determining Lyapunov Exponents From a Time Series", *Physica D*, **16**, 285-317.

Woodford, M. (1989) "Imperfect financial intermediation aand complex dynamics," *in Economic Complexity: Chaos, Sunspots, Bubbles and Nonlinearity*, ed. W.A. Barnett, J. Geweke, K. Shell, Cambridge University Press.

Wright, J. (1984) "Method for calculating a Lyapunov Exponent," *Physics Review*, **A29**, 2923-2931.

Yang, S-R. and Brorsen, B.W. (1992) "Nonlinear dynamics of daily cash prices," *American Journal of Agricultural Economics*, **74**, 706-715.

Yang, S-R. and Brorsen, B.W. (1993) "Nonlinear dynamics of daily futures prices: Conditional heteroscedasticity or chaos?", *The Journal of Futures Markets*, **13**, pp. 175-191.

Zarowin, P. (1988) "Size, seasonality and stock market overreaction," Manuscript, New York University, New York.

APPENDIX

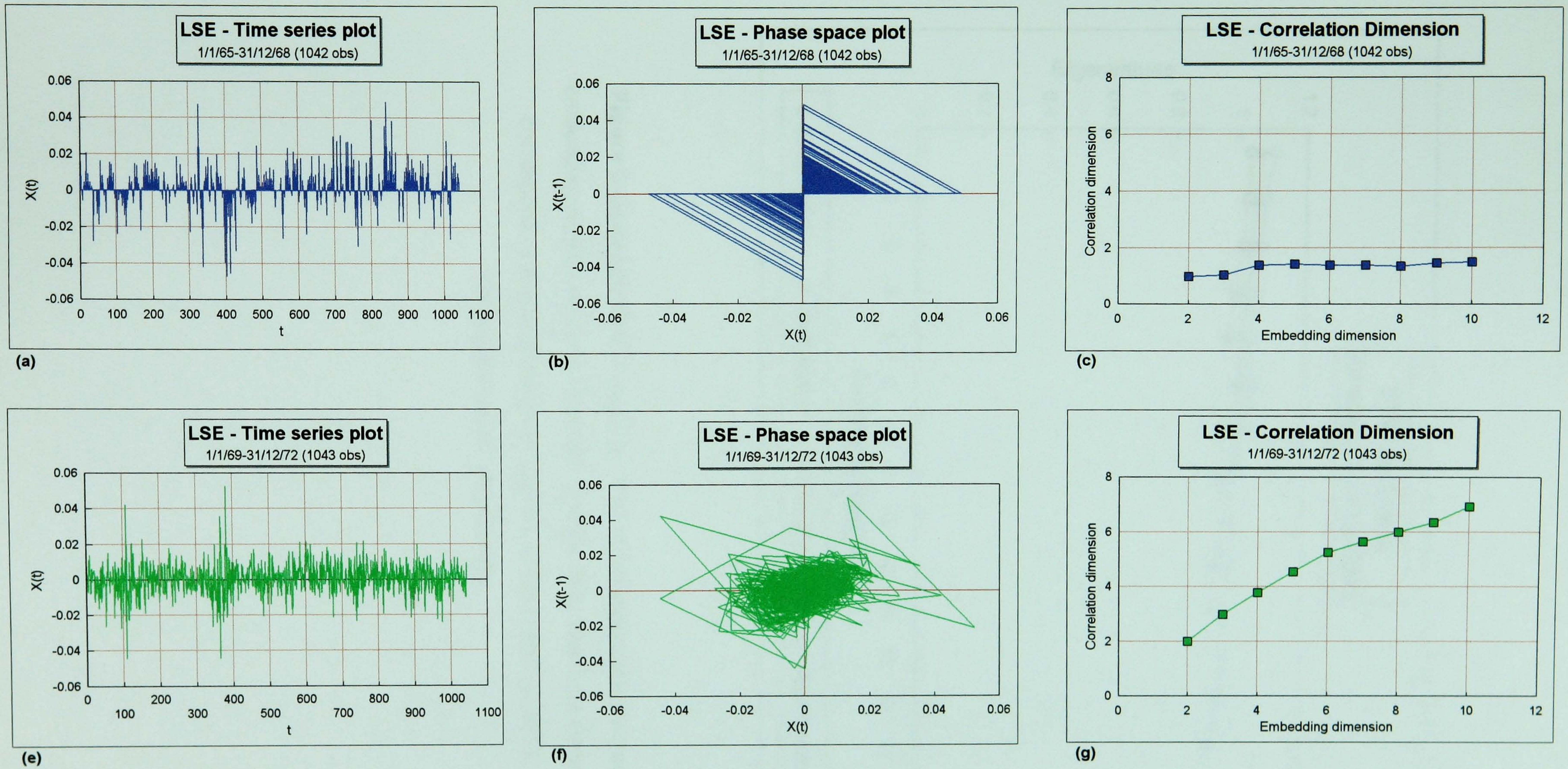

Figure 1 (appendix) : a-c : Time series plot, phase space plot and correlation dimension plot for the first 1042 obs. of our initial LSE data. e-g : The same plots for the next 1043 obs. of the same data.

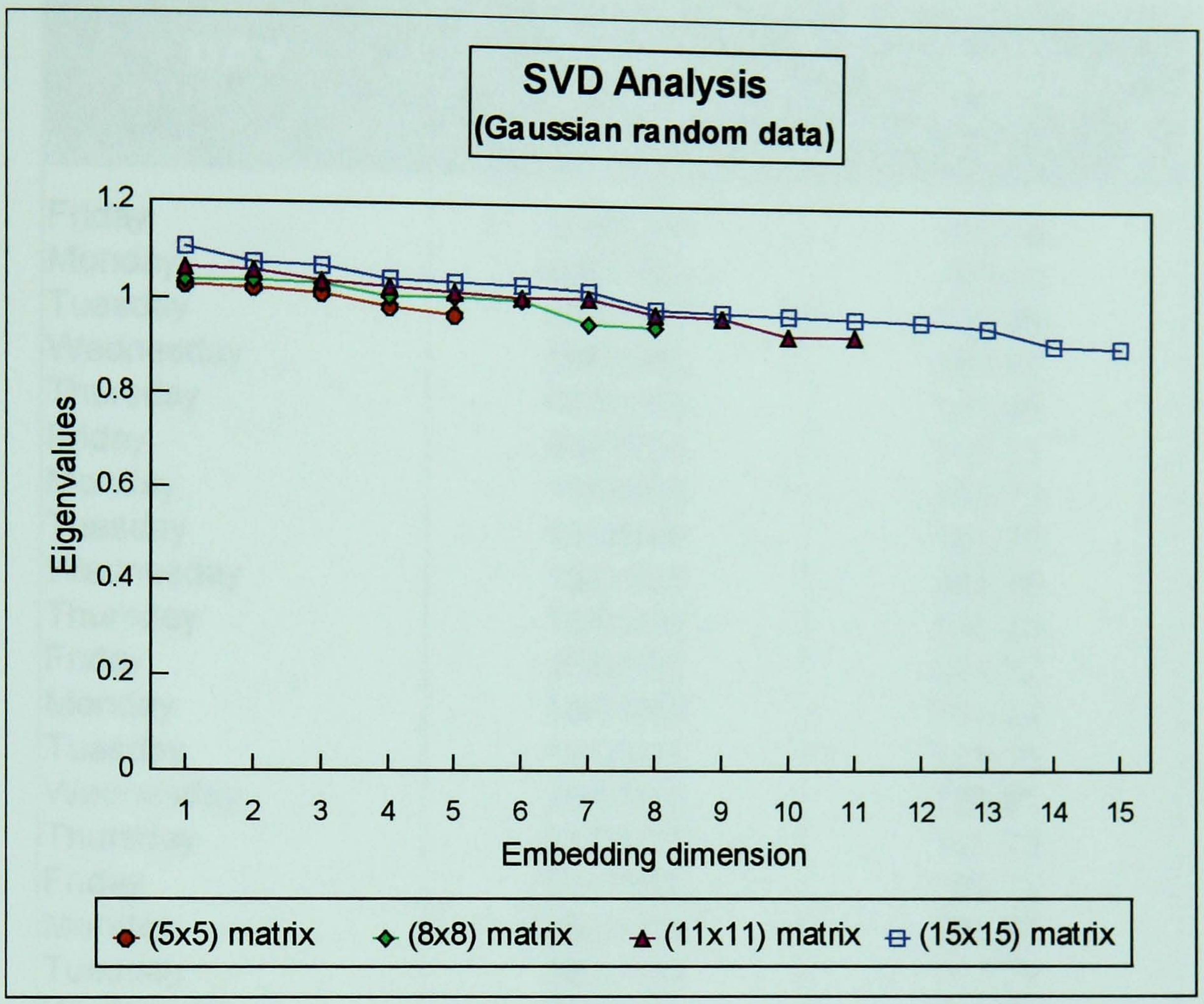

Figure 2 (appendix) SVD analysis - Plot of the eigenvalues versus embedding dimension for a Gaussian random data. The three curves correspond to diagonalization of a (8x8), (11x11) and (15x15) covariance matrix respectively .

TABLE 1 (APPENDIX)

LSE Index price from 1/1/65 to 26/2/65

DAY	DATE	INDEX
Friday	01/01/65	100.04
Monday	04/01/65	100.06
Tuesday	05/01/65	100.08
Wednesday	06/01/65	**101.67**
Thursday	07/01/65	101.69
Friday	08/01/65	101.71
Monday	11/01/65	101.73
Tuesday	12/01/65	101.75
Wednesday	13/01/65	**101.18**
Thursday	14/01/65	101.20
Friday	15/01/65	101.22
Monday	18/01/65	101.24
Tuesday	19/01/65	101.26
Wednesday	20/01/65	**101.71**
Thursday	21/01/65	101.73
Friday	22/01/65	101.75
Monday	25/01/65	101.77
Tuesday	26/01/65	101.79
Wednesday	27/01/65	**103.89**
Thursday	28/01/65	103.91
Friday	29/01/65	103.93
Monday	01/02/65	103.95
Tuesday	02/02/65	103.97
Wednesday	03/02/65	**104.92**
Thursday	04/02/65	104.94
Friday	05/02/65	104.96
Monday	08/02/65	104.98
Tuesday	09/02/65	105.00
Wednesday	10/02/65	**105.46**
Thursday	11/02/65	105.48
Friday	12/02/65	105.50
Monday	15/02/65	105.52
Tuesday	16/02/65	105.54
Wednesday	17/02/65	**105.76**
Thursday	18/02/65	105.78
Friday	19/02/65	105.80
Monday	22/02/65	105.82
Tuesday	23/02/65	105.84
Wednesday	24/02/65	**102.91**
Thursday	25/02/65	102.93
Friday	26/02/65	102.95

LIST OF ABBREVIATIONS

APT	Arbitrage Pricing Theory
AR	Autoregressive model
ASE	Athens Stock Exchange
BDS	Brock, Dechert, Scheinkman
B-E	Break Even
BM	Brownian Motion
CAPM	Capital Assets Pricing Model
CC	Correlation Coefficient
CMH	Coherent Market Hypothesis
DVS	Deterministic Versus Stochastic
EMH	Efficient Market Hypothesis
FBM	Fractional Brownian Motion
FD	Fractal Dimension
IID	Independently and Identically Distributed
LE	Lyapunov Exponents
LLE	Largest Lyapunov Exponent
LM	Logistic map
LSE	London Stock Exchange
MA	Moving Average model
MAE	Mean Absolute Error
MH	Maintained Hypothesis
MLP	Multi-Layer Perceptron
MOD	Method of Delays
MSE	Mean Squared Error
NN	Nearest Neighbours
NRMSE	Normalised Root Mean Squared Error
OBS	Observations
ODE	Ordinary Differential Equations
OLS	Ordinary Least Squares
OPT	Options Pricing Theory

PCA	Principal Components Analysis
PW	Piecewise Approximation model
R/S	Rescaled Range
RBF	Radial Basis Functions
RW	Random Walk model
SCF	Sign Change Frequency
SDIC	Sensitive Dependence on Initial Conditions
SSA	Singular Spectrum Analysis
STD	Standard Deviation
SVD	Singular Value Decomposition
SX	Simplex model